Space Syntax

Selected papers by Bill Hillier

Edited by Laura Vaughan, John Peponis and Ruth Conroy Dalton

UCL PRESS

Space Syntax

Space Syntax

Selected papers by Bill Hillier

Edited by Laura Vaughan, John Peponis and Ruth Conroy Dalton

UCL PRESS

First published in 2025 by
UCL Press
University College London
Gower Street
London WC1E 6BT

Available to download free: www.uclpress.co.uk

A CIP catalogue record for this book is available from The British Library.

ISBN: 978-1-80008-769-9 (Hbk.)
ISBN: 978-1-80008-770-5 (Pbk.)
ISBN: 978-1-80008-771-2 (PDF)
ISBN: 978-1-80008-772-9 (epub)
DOI: https://doi.org/10.14324/111.9781800087712

Contents

List of figures

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Chapter 9

Chapter 10

Chapter 11

Chapter 12

Chapter 13

Chapter 14

Chapter 15

Chapter 16

Chapter 17

Chapter 18

Chapter 19

Chapter 20

List of tables

Chapter 4

Chapter 11

Chapter 13

Chapter 14

Chapter 17

Chapter 19

Chapter 20

Note on the figures in this book

The compilation of the figures from the 20 pieces included in this volume was a vast exercise. We decided early on that where possible we would redraw images both to improve the readability of the many illustrations that have lost quality in the publication process, but also to introduce a uniform look to the book. There are some instances where we have preferred to use images from Bill Hillier's lectures, especially if these were in colour. Copyright has already been obtained for published papers, with third-party images updated with fresh copyright permissions. In a handful of instances we have been unable to obtain permission and have omitted the figure, but these are rare, and quite insignificant instances. The endnotes point to any significant changes or omissions, while the text has been revised to reflect two types of changes: where lengthy captions have been incorporated into the main body of the paper, or where colour images have replaced greyscale ones.

The editors are grateful to Emad Alyedreessy for his contribution to the book in redrawing many of its images. Dr Alyedreessy is a UK registered Architect who recently completed a PhD in Architecture at Northumbria University. His doctoral research project examined the relationship between spatial configurations and social solidarity in large-scale contemporary co-living buildings using several space syntax methods. The editors also wish to acknowledge the contributions to the illustrations made by Ruth Conroy Dalton, who created and refined many of the images.

The editors would like to give a special thanks to Dr Sheep Dalton for writing a new piece of software, *Phoenix,* which was used to reprocess old Axman software files to obtain production-quality images for the book.

Image credits

By Ruth Conroy Dalton: **15**. Specifically architectural theory: A partial account of the ascent from building as cultural transmission to architecture as theoretical concretion, Figure 15.40. All remaining line drawings were a joint effort by Dr Emad Alyedreessy and Professor Ruth Conroy Dalton, with the exception of the colour diagrams extracted from Bill Hillier's archives.

Reproduction rights for the 20 texts

The editors have the permission of the estate of Bill Hillier to reproduce all the following publications. We are grateful to Sheila Hillier for granting us these rights. All co-authors have provided permission for reproduction too. Following are the lists of publisher's permissions granted. We are grateful to all of them for doing so gratis. The editors have made every effort to obtain permission from the original publishers, though have not managed to do so in the case of a handful of publications. Following is a record of copyright permissions given.

Chapter 1 Hillier, B. (and Leaman, A.), with J. Musgrove and P. O'Sullivan. 1972. 'Knowledge and design'. In: Mitchell, W. (ed.), *AIA-Architect-Researchers' Conference. Environmental Design: Research and Practice. Proceedings of the Third Conference of the Environmental Design Research Association*. Los Angeles: University of California Press, 245–264.

Publisher: Environmental Design Research Association (EDRA): https://www.edra.org/. Reproduction granted via copyright information at: https://cdn.ymaws.com/www.edra.org/resource/resmgr/proceedings/toc/edra03-contents.pdf

Chapter 2 Hillier, B. and Leaman, A. 1973. 'The man–environment paradigm and its paradoxes'. *Architectural Design*, 43, 507–511.

Used with permission of John Wiley and Sons Ltd. from *Architectural Design*, 43, 1973; permission conveyed through Copyright Clearance Centre, Inc. on 22 August 2022.

Chapter 3 Hillier, B. and Leaman, A. 1974. 'How is design possible?' *Journal of Architectural Research*, 3, 4–11.

Published by Locke Science/AIA Press: untraceable. Subsequent owners Wiley cannot confirm copyright.

Chapter 4 Hillier, B. and Leaman, A. 1975. 'The architecture of architecture: Foundations of a mathematical theory of artificial space'. In: Hawkes, D. (ed.), *Models and systems in architecture and building.* Hornby: Construction Press, 5–28.

Published by Construction Press Ltd: untraceable.

Chapter 5 Hillier, B. and Leaman, A. 1976. 'Architecture as a discipline'. *Journal of Architectural Research*, 5, 28–32.

Published by Locke Science/AIA Press: untraceable. Subsequent owners Wiley cannot confirm ownership of the copyright.

Chapter 6 Hillier, B., Leaman, A., Stansall, P. and Bedford, M. 1976. 'Space syntax'. *Environment and Planning B: Planning and Design*, 3, 147–185.

Published by Sage. Reproduced here under authors' rights. See further information at https://uk.sagepub.com/en-gb/eur/pre-approved-permission-requests-journals.

Chapter 7 Hillier, B., Hanson, J., Peponis, J., Hudson, J. and Burdett, R. 1983. 'Space syntax: A different urban perspective'. *The Architects' Journal*, 178(48), 47–63.

Published by *The Architects' Journal*. No response to three requests.

Chapter 8 Hillier, B., Hanson, J. and Peponis, J. 1984. 'What do we mean by building function?' In: Powell, J., Cooper, I. and Lera, S. (eds.), *Designing for building utilisation*. London: Spon, 61–98.

Published by Spon Press (part of Taylor & Francis). Publisher unable to grant copyright and has no record of book. Copyright on this paper belongs to the authors.

Chapter 9 Hillier, B. 1985. 'Quite unlike the pleasures of scratching: Theory and meaning in architectural form'. *9H*, 7, 66–72.

Published by UCL Bartlett School of Architecture and Planning. Published under author's rights.

Chapter 10 Hillier, B. 1985. 'The nature of the artificial'. *Geoforum*, 16, 163–178.

Used with permission of Elsevier Science & Technology Journals from *Geoforum*, 16, 1985; permission conveyed through Copyright Clearance Centre, Inc. on 25 August 2022.

Chapter 11 Hillier, B., Hanson, J. and Graham, H. 1987. 'Ideas are in things: An application of the space syntax method to discovering house genotypes'. *Environment and Planning B: Planning and Design*, 14, 363–385.

Published by Sage. Reproduced here under authors' rights. See further information at https://uk.sagepub.com/en-gb/eur/pre-approved-permission-requests-journals.

Chapter 12 Hillier, B. 1988. 'Against enclosure'. In: Teymur, N., Markus, T. and Wooley, T. (eds), *Rehumanizing housing*. London: Butterworth, 63–88.

Used with permission of Elsevier Science & Technology Journals from *Rehumanizing housing*, 1988; permission conveyed through Copyright Clearance Centre, Inc. on 21 August 2022.

Chapter 13 Hillier, B. 1989. 'The architecture of the urban object'. In 'Space syntax: Social implications of urban layouts'. Special issue, *Ekistics*, 56 (334–5), 5–21.

Published by *Ekistics* journal. Permission granted by the publisher on 12 April 2022.

Chapter 14 Hillier, B., Penn, A., Hanson, J., Grajewski, T. and Xu, J. 1993. 'Natural movement: Or, configuration and attraction in urban pedestrian movement'. *Environment and Planning B: Planning and Design*, 20, 29–66.

Published by Sage. Reproduced here under author's rights. See further information at https://uk.sagepub.com/en-gb/eur/pre-approved-permission-requests-journals.

Chapter 15 Hillier, B. 1993. 'Specifically architectural theory: A partial account of the ascent from building as cultural transmission to architecture as theoretical concretion'. *Harvard Architecture Review*, 9, 8–27.

Published by Princeton Architectural Press. Publisher cannot confirm ownership of the copyright.

Chapter 16 Hillier, B. and Penn, A. 1994. 'Virtuous circles, building sciences and the science of buildings: Using computers to integrate product and process in the built environment'. *Design Studies*, 15, 332–365.

Used with permission of Elsevier Science & Technology Journals from *Design Studies*, 15, 1994; permission conveyed through Copyright Clearance Centre, Inc. on 25 August 2022.

Chapter 17 Hillier, B. 2002. 'A theory of the city as object: Or, how spatial laws mediate the social construction of urban space'. *URBAN DESIGN International*, 3–4, 153–179.

Reprinted by permission from Springer Nature. 'A theory of the city as object: Or, how spatial laws mediate the social construction of urban space' by B. Hillier Copyright 2002.

Chapter 18 Hillier, B. and Netto, V. 2002. 'Society seen through the prism of space: Outline of a theory of society and space'. *URBAN DESIGN International,* 7, 181–203.

Reprinted by permission from Springer Nature. 'Society seen through the prism of space: Outline of a theory of society and space' by B. Hillier and V. Netto Copyright 2002.

Chapter 19 Hillier, B. 2012. 'Studying cities to learn about minds: Some possible implications of space syntax for spatial cognition'. *Environment and Planning B: Planning and Design*, 39, 12–32.

Published by Sage. Reproduced here under author's rights. See further information at https://uk.sagepub.com/en-gb/eur/pre-approved-permission-requests-journals.

Chapter 20 Hillier, B. 2012. 'The city as a socio-technical system: A spatial reformulation in the light of the levels problem and the parallel problem'. In: Müller Arisona, S., Aschwanden, G., Halatsch, J. and Wonka, P. (eds), *Digital urban modelling and simulation.* Heidelberg, Dordrecht, London and New York: Springer, 24–48.

Used with permission of Springer Nature, from *Digital urban modelling and simulation*; permission conveyed through Copyright Clearance Centre, Inc. on 25 July 2023.

Foreword

The publication of this book coincides with the 50th anniversary of the programme of architectural research known as space syntax, which was initiated by Bill Hillier and colleagues in the early 1970s. While his earliest work took place at the Royal Institute of British Architects, Hillier was appointed director of the Unit for Architectural Studies in the then Bartlett School of Architecture and Planning at University College London in 1974, the year when one of his pioneering papers with Adrian Leaman, 'How is design possible?', was published. The first article on the research programme of space syntax, co-authored with Adrian Leaman, Paul Stansall and Michael Bedford, followed in 1976 – with a series of essential papers developing the theories emerging in the following years – all of which appear, in freshly illustrated transcriptions, in this volume.

A comprehensive theory for the discipline of space syntax was first presented in *The social logic of space*, co-authored by Bill Hillier with one of his most important collaborators, Julienne Hanson, in 1984. The influence of the discipline beyond academe has been profound from the start, while the first such project to attract public attention was the masterplan for the King's Cross area, developed by Norman Foster in 1988. With the growing demand from policy and practice for input from the theories and methods, Space Syntax Ltd was formed in 1989 to focus on the application of space syntax in professional practice through design, policy development or performance modelling.

Another landmark came when Hillier reported the latest fundamental insights from research in *Space is the machine*, published in 1996, where he revisited the underlying theoretical premises of space syntax. A year later, the first international symposium on space syntax was held in London, attracting a wide variety of papers. There have been 13 more biennial symposia since then, held in Brasília (1999), Atlanta (2001), London (2003), Delft (2005), Istanbul (2007), Stockholm (2009), Santiago (2012), Seoul (2013), London (2015), Lisbon (2017), Beijing (2019), Bergen (2022) and Nicosia (2024). Space syntax research is now pursued in many universities and countries around the world and space syntax literature, both scientific papers – the Web of Science identifies about three thousand papers when queried with keyword 'space syntax' – and scholarly books is burgeoning. It is likely that more of the ideas, concepts and techniques of analysis originally developed in the context of space syntax will continue to be integrated into the mainstream discourse in architecture, urban design and planning.

Few programmes of architectural research have benefitted from a trajectory of continuous and consistent development and application in professional practice over such a prolonged period of time. This book, however, does not aim to chart the development of the field of space syntax as a whole. Rather, it highlights the evolution of the ideas of Bill Hillier by bringing together a selection of his papers published between 1974 and 2012, whose ideas remain fundamental to space syntax as a programme of research. At the same time, as indicated by the growing citation record of Hillier's major books, they are of interest not only to many of the people who study architecture as a discipline but also to those working in fields of enquiry as diverse as archaeology, anthropology, sociology, environmental psychology, cognition, geography and computational spatial analysis.

Each of the papers in this collection is introduced in a short essay by a different scholar, offering a perspective on its significance, a critical commentary and, in some instances, suggestions regarding extensions yet to be explored. Thus, the essays are intended to prompt the reader's own critical reflection. The Introduction provides an underpinning to the ideas presented throughout this book. The original texts are reproduced with the minimum corrections necessary.

Acknowledgements

This mammoth effort would have been impossible without the intellectual contributions of our colleagues from across the space syntax community over the past 50 years. In addition, we would like to thank our editors at UCL Press, and in particular:

UCL Library's Bartlett, Stores, and Open Access teams who went above and beyond in obtaining copies of the most obscure publications to enrich the bibliography of this book.

Copyeditor Tim Rutherford-Johnson, for his incredible eye for detail down to every last formula, caption and footnote.

Julie Willis and her team at Westchester Publishing Services, for their endless care with image reproduction and graphic layout.

Dr Garyfalia Palaiologou for her expertly prepared index.

Lastly, to Sheila Hillier and the family for their support of our producing this book, which we believe Bill would have loved to have seen.

Introduction

Keys to Bill Hillier's thought on architecture

John Peponis, Ruth Conroy Dalton and Laura Vaughan

An introduction recognising, tracking and summarising all the multiple strands of ideas developed by Bill Hillier over the course of 40 years would compete with the main aim of this book, which is to provide the reader with direct access to original papers. Significant effort would also be required to situate Hillier's ideas within the context of the extensive and diverse bodies of literature to which he refers, or with which he is known to have engaged, at least by his students and close associates. Our first and more limited aim here is to identify key propositions and theses to help readers make their own maps and connections. This will be followed by brief comments on some of the threads and themes that run through the work.

Buildings have four generic functions

(Hillier and Leaman 1972; 1974; 1975; 1976; Hillier, Musgrove and O'Sullivan 1972)

Buildings mediate social relationships between people and also between people and nature. They are defined according to four generic functions. They modify activity and behaviour, climate and environment, material and economic resources, and the symbolic systems that constitute a culture. The effects of buildings upon the activities and behaviours of their inhabitants are mediated by cultural norms, conventions, habits and ways of thinking. The effects upon the environment are mediated by the deployment of technological and human resources. Thus, building functions can be studied from two points of view: the performance of a given building or set of buildings; and the collective knowledge and languages of description and representation that underpin the design of families of buildings for human purposes at a particular point in time. At their best, architectural research and theory should be organised into different strands according to the general functions addressed.

The idea of generic function is fundamental to the growth of Hillier's research programme and to the growth of architectural research at large. Generic functions provide an entryway and the keys to the analysis of spatial form. Lawfulness in morphological phenomena is unlikely to be revealed when we address specific itemised needs or programmatic requirements. These are built upon generic affordances and functions and the task is to understand how this is possible and the forms it takes. In Hillier's own work, of course, the focus is upon analysing how buildings modify activity, behaviour and the symbolic systems that constitute culture; climate and resource modification are not discussed.

Design is made possible by cognitive pre-structures

(Hillier and Leaman 1974; 1975)

A simple enclosure, one of the most elementary moves by which space is organised for human purposes, is already a structure of relationships. An interior is distinguished and set apart from the circumjacent space. By implication, a surrounding neighbourhood is defined whose outer boundary needs not be demarcated; what happens in this neighbourhood, for example whether another enclosure can be built, or an activity occur, becomes associated with the initial enclosure by virtue of spatial proximity. For the interior to be useful, an opening is required which simultaneously connects the interior to the exterior and allows the exercise of control over the threshold. Finally, the possible creation of another enclosure, in the neighbourhood, or at a distant and independently defined location, entails a potential relationship of similarity whereby multiple interiors are recognised as members of a set, or class, distributed across space. Thus, the potential for categoric differentiation and definition, the potential for control over the relationship of differentiated categories, and the potential for relationships based on spatial proximity or structural similarity across space are all intrinsic to the creation of an elementary enclosure. Every new physical intervention – every new wall, enclosure, or opening – has an effect on the spatial relationships already present, as will be discussed later.

The encoding of logical relationships into spatial form provides a basis for built space to be organised and to function socially. In this way, a close relationship is established between the physical world and our mental conception of it. As design operates on physical entities to create spatial relationships, so it also reproduces or transforms the way in which space is conceived. Codes linking the organisation of physical boundaries to logical relationships and to generic functions are the deeper cognitive pre-structures that make design possible. The task of theory is to bring such codes to light and to support their intentional transformation.

Design is knowledge-based creativity

(Hillier, Musgrove and O'Sullivan 1972; Hillier 1996)

Design engages different kinds of knowledge. First, knowledge of precedent, including prevalent solution-types to recurring and familiar design problems. Second, knowledge of tools or 'instrumental sets', including available technologies, materials and types of construction. Third, critical reflection and theory-making regarding the underlying principles and assumptions, often tacit, through which form and its functions are understood. This includes the development of testable models of building performance relative to the generic functions.

The relation of design to knowledge is best approached in terms of testing conjectures. Conjectures, at the earlier stages of design, spring from a designer's prior knowledge of design possibilities and appropriate responses. Analytical theories, systematic comparative analyses of precedents and research-based performance models come into play as these conjectures are tested to be rejected, transformed, developed or accepted. The conjecture-test model of design embraces the dialogue between imagination and knowledge.

Insofar as architecture and the formulation of architectural design intent involve a critical questioning of solution types, instrumental sets, and underlying assumptions about the relationship of form and function taken for granted in the production of buildings, one can say that architecture 'adds theory' to building.

Socio-spatial morphology as a function of topological restrictions on a random process of growth

(Hillier and Leaman 1975; Hillier, Leaman, Stansall and Bedford 1976; Hillier and Hanson 1984)

Morphology is the systematic study of possible forms based on explicitly formulated generative principles. In Hillier's early work, spatial morphology is founded on the idea of restrictions placed upon an otherwise random process of growth. Growth is conceived as a pattern of aggregation and arrangement of elementary enclosures, such as described above. In the absence of restrictions that have a social reasoning, new enclosures might be scattered randomly over the land, subject only to limitations arising from geography and landscape. The question then becomes: what form might potential restrictions on such a random process take? In the foundational texts, restrictions are formulated in terms of basic topological relationships, such as those that Piaget argued are at the foundation of the development of human spatial cognition.[1] Elementary structures can be adjacent and contiguous to each other, inside each other, or placed so as to define an open space between them. Adjacency, insidedness, betweenness and permeability – which are necessary to provide access into each enclosure or between contiguous enclosures – are, therefore, the fundamental relationships.

In Hillier and Leaman (1975) the variability of the structures arising from the application of different aggregation rules is discussed with an emphasis on the density or sparseness implied by the restrictions placed by different generative rules. By the time of the publication of 'Space syntax' (Hillier, Leaman, Stansall and Bedford 1976), only a year later, eight possible generative rules had been identified that, working severally or in conjunction, were able to account for the rich variety of recorded forms across cultures and historical periods. To facilitate the precise definition of the generative rules, a notational language was being developed in parallel, right from the outset. This found its most complete and concise formulation in *The social logic of space* (Hillier and Hanson 1984), which presents the eight generative rules previously identified in 'Space syntax' with some refinements of definition and a new interpretation of their underlying logic.

The underlying logic of the generative rules is fundamental to the conceptual fabric of early space syntax theory. One dimension of logical variability is the extent to which the spatial structure that emerges from a process of aggregation is defined equally by all elementary units involved, or dominated by only a few of them. For example, when elementary enclosures are arranged to surround, and thus define, an open space between them, the emergent structure of containment (of a courtyard or other common ground) is *distributed*: all units partake equally in its definition. When, however, units are aggregated inside a larger primary enclosure, the emergent structure is *nondistributed*: the surrounding boundary plays the primary role in restricting the future relationship between the units. The other dimension of logical variability is the extent to which elementary units are syntactically (space syntax commonly uses the term 'syntactic' and its derivatives) interchangeable or not. When, for example, units are aggregated according to a simple rule of adjacency, their relationship is *symmetrical*. Each unit equally participates in the emergent pattern and thus can be interchanged with others. When, however, each new enclosure is placed around another, the relationship between the elementary enclosures is *asymmetrical*. Each holds a unique place in the order of containment.

As important to the logical relationships between generative rules is the distinction between the rules themselves and their outcomes. Once restrictions are placed on a process of growth or aggregation which is otherwise random, laws of a mathematical nature continue to operate and to affect the outcome. Outcomes can thereby have systematic properties that are not explicitly inscribed in the rules themselves. Take, for example, the process of aggregation of paired couplets of an enclosure and an open space according to the rule that the open space of the next couplet should be joined to the open space of a couplet already present in the pattern. The rule itself only specifies a local adjacency. But the emergent pattern has additional properties: the network of open spaces is continuous and forms circulation loops around building-islands of varying sizes; it is everywhere characterised by the succession of narrower and wider open spaces; and, in the earlier stages of growth, every open space in the middle of the aggregate is linked to the outer perimeter by an unobstructed straight path. In short, the emergent global structure has its own characteristic properties which are not inscribed in the rule itself. Once these properties are understood, human agency can intervene to direct further evolution. For instance, the unobstructed links between the innermost open spaces and the perimeter can be preserved over future growth by additional restrictions, thus giving rise to continuous linear successions of open spaces, akin to irregularly delimited 'main streets' – a possibility taken up by Hillier in 'The architecture of the urban object' (1989).

The logic of the morphogenetic model described is already social, in the same sense that language is already social: the model rests on the cognitive capacities of the human mind; its structure exists independent of the individual subjects that may deploy it; and its application occurs in the context of a social organisation of human activity. But how can spatial forms be associated with different kinds of social organisation? In an earlier work (Hillier and Leaman 1975), two interlinked ideas are brought to bear on the question. First, that dense agglomerations indicate an instrumental use of space, to facilitate encounters, interactions and exchanges by bringing things together; sparse aggregations, on the other hand, indicate that space is used to express symbolic orders, to keep spatially demarcated categories apart. Second, that dense agglomerations are associated with the form of social solidarity that Durkheim (1933) calls organic: society is held together by the division of labour leading to interdependence. Sparse agglomerations are associated with a form of solidarity that Durkheim calls mechanical: society is held together by common beliefs, rituals and customs, a shared symbolic order. Perhaps the two ideas can be expressed very simply: spatial boundaries can be used to include and to exclude; this is the foundation of the social logic of space. Sometimes, space works to create relationships that integrate the differentiated parts of society. Sometimes it is used to keep the differentiated parts apart. Sometimes space works to reinforce the primary principles of social solidarity. Sometimes it works to provide an alternative basis for social encounters and relationships (Hillier, Leaman, Stansall and Bedford 1976). Thus, space does not 'mirror' society but acts as one of its constitutive dimensions.

Representing and measuring spatial relationships: From organising principles to perceptual affordances

(Hillier, Hanson, Peponis, Burdett and Hudson 1983; Hillier and Hanson 1984; Hillier, Burdett, Peponis and Penn 1987)

The measurement of spatial relationships for the sake of comparative analyses was initially consistent with the focus of socio-spatial morphology on topological relations. Imagine a building with all its doors closed. The interior is divided into a number of discrete spaces, each with its enclosing boundary, however irregular. Open the doors and identify the permeable connections between the constituent spaces. Represent the relationships thus identified as a graph, where each node represents a space and each edge represents a permeable connection.[2] The distance between any two spaces is measured by the minimum number of thresholds, or intervening spaces that must be crossed to get from one to the other. Hillier and Hanson (1984) use the term 'depth' to describe distance measured in this way.

The graph representation captures the underlying topology. Four basic unweighted graph-theoretic measures are applied, to describe each space. The number of direct connections to neighbouring spaces is a measure of local *connectivity*. The sum of the reciprocals of the number of connections of each neighbouring space is a measure of how far each space exercises *control* over local access to its neighbours. The mean depth from all other spaces is a measure of 'closeness centrality'; Hillier and Hanson use the term *integration*. The proportion of shortest paths between all possible pairs of origins and destinations that go through a space is a measure of 'betweenness centrality'; Hillier and Hanson use the term *choice*. Each space in an interior is described according to these measures and also by its depth from the 'carrier' – a term used to describe circumjacent space external to the system.

Closeness centrality is linked to the idea of asymmetry. The greater the boundary-distance between spaces, the greater the asymmetry in that access to some spaces depends on first traversing others. On the other hand, the presence of circulation loops implies that control over access is distributed: there are alternative ways to reach spaces. This is captured by lower choice values. Thus, the unweighted graph-theoretic measures of the properties of spatial patterns are congruent with the underlying dimensions of variability of the generative rules.

The ideas described above cannot be applied to continuously flowing space, whether building open plans or irregular rooms, or open public spaces and street systems. The resolution of the problem of analysing continuously flowing space is a major contribution of the collaboration of Hillier with Hanson. It is proposed that continuous space can be broken up into discrete connected spatial elements in two ways. The relevant *locally* circumscribed unit is a convex space, that is, a space any two points of which can be linked by a straight-line segment that lies entirely within the space. A layout is partitioned down to the fewest and 'fattest' (namely, approximating a simple square or circle) connected convex spaces that could fully cover it – an intuitive rule that has not been amenable to exact algorithmic interpretation. On the other hand, the relevant *extensive* unit is the uninterrupted straight line of sight that constitutes potential movement along an axis. A layout is covered by the fewest and longest continuously intersecting lines (the term used was 'axial lines') that would cover all the convex spaces and all the possible circulation paths around boundaries. A precise algorithmic application of this rule was proposed years later,[3] and efficiently implemented, with some original modifications, by Turner and colleagues.[4]

To arrive at calculations of centrality, distances were intentionally measured in terms other than path length. In the case of the line-based map, the measure of distance was associated with direction changes occurring as one moved from one line to another. In the case of the convex map, distances were measured by the number of convex spaces that had to be crossed to go from one place to another.

The invention of the convex-based and the line-based representations of spatial patterns carried with it some fundamental shifts in the intellectual foundations of morphological analysis, relative to the earlier period. First and foremost, analysis became sensitive to dimension and shape. The number of straight lines needed to represent a curved street depends on the street's width and the street's curvature; similarly, two slightly offset linear streets meeting the two sides of a main street at two proximate but distinct T-junctions could be represented by a single line cutting diagonally through them, or require separate lines depending on street width and the distance between the offset intersections. Thus, while the measures applied in analysis were still not directly responsive to metric shape or dimension, the underlying representations upon which the measures were based became shape- and dimension-sensitive.

Second, the new representations were more tightly linked to perception and the affordances of the environment relative to basic human functions of space. A convex space, for example, allows each of its potential occupants to see and address all others, a relation of generalised reciprocity that would not prevail in an L-shaped room. A long line of sight and potential movement extending out from a particular location provides orientation relative to potential destinations. It can also provide the opportunity to pass by any number of individual building entrances along the way, thus linking them by association with a linearly extended visual horizon.

The analysis of spatial patterns in terms of relationships that underpin perceptual affordances was further enriched when visibility polygons, or 'isovists',[5] were added to line-based and convex-based representations.[6] A visibility polygon represents, in plan, all the area that is readily visible and accessible from a point. Its perimeter is defined by two alternative conditions: physical boundaries that manifestly block views and movement; and occluding edges, that is, edges of the visual field that are open towards further accessible space 'around the corner'. The family resemblance of ideas that made the incorporation of visibility polygons into Hillier's thinking natural is easy to intuit. Walking to the occluding edge of a visibility polygon to then change direction and move across it into previously invisible extensions of space is analogous to reaching towards the end of a straight line of movement to then change direction and move along the next intersecting line. One is essentially dealing with the way in which continuously linked space unfolds around boundaries as one moves in it.

The sensitivity of spatial analysis to geometric properties was further enhanced when angular distance was substituted for lines distance thanks to the conceptual innovations and computational solutions introduced by Dalton and by Turner.[7] Lines distance does not take into account the angles involved in direction changes. It is based on the minimum number of turns required to go from one place to another irrespective of the rotation of each line relative to another. Angular distance is measured by the minimum aggregate rotation angle required to reach one place from another. In the later papers by Hillier (2012a, 2012b, 2016), both closeness centrality (integration) and betweenness centrality (choice) are regularly measured based on angular distances as well as turns distances.

Space as a morphic language and 'space syntax'

(Hillier, Leaman, Stansall and Bedford 1976; Hillier, Hanson, Peponis, Burdett and Hudson 1983; Hillier and Hanson 1984)

Linguistic meaning is encoded in the social creation of a vocabulary, but also through the arrangement of words in sentences that meet grammatical and syntactic rules. While natural languages have very rich vocabularies, artificial languages such as mathematics and musical notation have limited vocabularies of symbols. In the case of mathematics, meaning resides in the creation and observance of the rules by which symbols are combined into valid statements. In the case of music, meaningfulness arises from the ways in which a small repertoire of notes and relative durations is deployed to create complex compositional patterns. Hillier and his colleagues, including at that time Paul Stansall, suggested that built space is a morphic language: it is based on the deployment of a small vocabulary (physical boundaries and separations, spatial demarcations and differentiations, accessible connections and visual connections) that is used to create complex spatial structures that are characteristic of cultures, organisations, institutions, communities and social groups. This insight has led to the adoption of the term 'space syntax' to describe the approach to spatial morphology which was developed to address the social logic and functions of built space.

'Space syntax' analysis came to encompass the earliest emphasis on abstract topological relationships and the later emphasis on relationships descriptive of generic perceptual affordances. The work of Hillier has typically not addressed aspects of spatial vocabulary commonly associated with different styles: for example, whether all corners of a room are built, as in the classical tradition, or whether a room is defined by free-standing walls that extend outwards from its interior without interruption, as is often the case with the open plans of Mies van der Rohe, or whether the corners of a room are left visually open, as is sometimes the case with Frank Lloyd Wright. Whether the idea of built space as a morphic language can be extended or refined to offer new insights on the ways in which underlying relationships are realised in particular design languages remains an open question in the work of Hillier.

Syntactic dimensions of the social categorisation of space

(Hillier and Hanson 1984; Hillier, Hanson and Peponis 1984; Hillier, Hanson and Graham 1987)

The emphasis on syntactic relationships brought forth a shift in the approach to the social categorisation of space. Earlier work (Hillier and Leaman 1975) envisaged the possible development of a theory of how particular constellations of ideas and distinctions used to describe types of space arise in conjunction with the evolution of the spatial form. Later work addressed different questions. First, given a set of socially differentiated and labelled spaces, what are the syntactic relationships between the spaces to which the labels are assigned? For example, in a sample of cottages and terraced houses in England, the living room, a space of everyday life and interaction, is always more integrated into the rest of the ground floor than the kitchen, a room traditionally dominated by women; the kitchen is in turn more integrated than the parlour, a 'best room' at the front of the house, typically reserved for special occasions (Hillier and Hanson 1984). The invariant order of integration of such commonly labelled domestic spaces is akin to a socio-spatial genotype. The key idea is that categorial differentiation is mapped onto syntactic differentiation: different labels, and hence functions and roles, are assigned to spaces which are syntactically non-equivalent and thus non-interchangeable. Spaces assigned different function labels, such as 'living room', 'classroom' or 'patient room', essentially enter syntactic relationships as elements of distinct kinds; thus, in a home, a school or a hospital, syntactic rules may apply to the relations of the spaces of the same label relative to the spaces of another: living rooms may be more integrated than kitchens, classrooms are likely to be grouped by corridor or wing and have easy access to libraries or resource centres, and patient rooms are likely to be visible from or at least easily accessible from nurses' stations.

Hillier allows that the assignment of different labels to previously equivalent positions adds restrictions to their interchangeability which are not evident from the mere visual inspection of a layout. For example, in the Bororo villages described by Lévi-Strauss, residential huts are arranged around a men's house. However, the apparent interchangeability of any hut with any other disappears when one takes into account that the space of the village is divided by two diametric oppositions: there are two exogamous and intermarrying moieties and the huts of each occupy the north and south halves of the circumference of the circle, respectively. Furthermore, the eight clans are split into an upper and lower group – or an upstream and a downstream group when the village is built near a river – and the huts of each group are also arranged on distinct halves of the circumference along the east–west axis. Thus, in some cases, such as the Bororo villages, the social categorisation of space is not grafted onto concomitant and invariant syntactic non-equivalence but adds semantic restrictions to the way in which otherwise equivalent spaces are assigned to occupants and uses. We note that Hillier and Hanson (1984, p. 206) use the expression 'semantic illusion' as part of their argument that the effects of semantics can be incorporated into more elaborate syntactic descriptions and thus do not challenge the underlying syntactic analytical framework.

The social logic of spatial syntaxes, however, revolves not only around categorisation but also around the manner in which generic syntactic properties are invested with social function over and above categorisation. For example, the most integrated spaces might be associated with main functions of everyday life; or they may be narrow corridors, passages or lobbies. In the first case, integration works to bring people together and create opportunities for interaction; in the second case, it works to maintain the separation of the main functional spaces, even as it facilitates transitions from one to another; particularly so when openings are so disposed as to restrict views from one main space to another, across the mediating passage.

The morphology of movement and co-presence and the creation of social life

(Hillier, Hanson and Peponis 1984; Hillier, Burdett, Peponis and Penn 1987; Hillier 1988; Hillier 1989; Hillier and Penn 1991; Hillier, Penn, Hanson, Grajewski and Xu 1993; Hillier 1996; Hillier and Tzortzi 2006)

The idea that the spatial organisation of the built environment enhances or limits everyday opportunities for co-presence and interaction has been present since the early work of Hillier and Leaman (1975). It remains fundamental to the theoretical models of the relationship between space and society advanced in *The social logic of space*. Indeed, Hillier and Hanson (1984) go beyond the question of whether space supports denser or sparser, inclusive or exclusive, patterns of co-presence and interaction. They set the foundations for an analysis of distinct types of encounter patterns treated as a spatio-temporal morphology, a direction of enquiry later pursued, but only partly, by others[8] taking a network analysis approach to the study of social relationships. Nevertheless, a critical moment in the development of the research programme of space syntax came with a subsequent and initially serendipitous discovery. In order to test the idea that housing estates were associated with lower rates of human presence in open public spaces compared to ordinary streets, rates of encounter were measured. An unexpected positive and strong correlation was found between the number of people moving in a particular street within the area of Barnsbury, in London, and the degree of integration of that street into the surrounding street network. The correlation was subsequently found to hold in different areas of London, at a variety of scales of spatial analysis, and became the foundation of the idea of 'natural movement'.

'Natural movement' was an expression used by Hillier to refer to the fact that the distribution of people over the streets of an urban area is determined by the syntax of the street network, even though the aggregate volume of people is obviously affected by land use and development density – the correlation discovered while studying areas in London was subsequently confirmed for samples of cities around the world. In essence, urban centrality describes both a morphological condition, as measured by integration, and a social outcome, as measured by the volumes of pedestrian movement and presence.

Several arguments were pursued in extension and in interpretation of this empirical finding. First, urban space was seen to support an ambient sense of community, distinct from the communities of interaction and the communities of common interest studied by sociologists – Hillier used the term 'virtual community' to describe the sense of being a member of a society that emerges from regular encounters with others, familiar or unfamiliar, recognised or anonymous, locals or visitors, in public space. Second, the effects of urban space on the distribution of movement and co-presence become a dimension of urban economy: retail and other land uses that depend upon the passing trade seek the advantage conferred by syntactic centrality. The extent to which syntactically central, or integrated, streets are distributed throughout the urban realm, or clustered in particular locations, becomes associated with the distribution of social and economic vitality at various scales.

The pattern of natural movement did not hold in studies of social housing estates, where the small number of people present in public space decreased as distance from the perimeter, measured by the number of turns, increased. This was interpreted as having resulted from various kinds of discontinuity that characterise the relationship of estates to the circumjacent street network. Entrances facing inward into courtyards; unoccupied open spaces functioning as boundaries; proximate and successive changes of direction discouraging entry for those not living on the estate; labyrinthine internal layouts: these were all forms of 'hard' or 'soft' discontinuity. Paradoxically, designs that were aimed at strengthening local communities were said to be functioning to create enclaves of isolation, by eliminating through-movement and by encouraging inhabitants to find the shortest ways out. Perhaps this resulted from an association, in the mind of some architects and sociologists, of local community with enclosure and territoriality. Design motivation aside, however, research showed that certain types of local layouts cause a breakdown of what otherwise appears as a generic association between the syntax of street networks and the distribution of movement and encounter. Supposing the creation of virtual community to be an important contribution of urban space to the creation of society, this means that estate designs were dysfunctional.

Movement densities and syntactic integration were also correlated in some building interiors, including work environments (Hillier and Penn 1991) and museums (Hillier and Tzortzi 2006). And yet, in contrast to street networks, building interiors are designed in response to complex and type-specific programmatic and functional requirements, associated with different kinds of organisational, managerial or institutional practices (Hillier, Hanson and Peponis 1984). Thus, the ways in which the syntax of space supports or hinders co-presence and encounter depends upon the interfaces between circulation and a rich variety of behavioural settings. Of course, the sense of liveliness of a street is linked to the interface between private interiors and public space, the number of entrances or the spread of shop frontages. In the case of building interiors, however, the idea of a movement interface needs to be extended to encompass the relationship between circulation and meeting rooms, formal and informal, large and small, enclosed or open; social amenities such as lounges, coffee bars or break areas; supporting facilities such as libraries, resource and information centres; reception lobbies, halls and waiting rooms; and, of course, the different work-related functions, such as, for example, desk-stations, laboratory benches, workshop areas or project rooms. In the face of the variety of the encounter-generating interfaces between movement and specific behavioural settings, one useful generalisation is the distinction between circulation spaces that are self-contained and well separated from main uses, as is often the case in large hospitals, and circulation spaces that are more open towards the circumjacent areas they traverse as is usually the case in retail environments or open-plan offices of different kinds.

The correlation between the distributions of movement and syntactic integration, and the idea that movement is a generator of encounter opportunities, became central tenets in Hillier's thinking about the social functions of space. They provided a basis for fleshing out the earlier hypothesis that the social function of architecture is not merely to frame social life, through patterns of categorisation and control over access, but also to generate it. The design of built space creates social life when it creates opportunities for co-presence, interaction and co-awareness.

Three kinds of lawfulness

(Hillier 1985)

The discovery of the natural movement principle led Hillier to define three kinds of lawfulness that underpin the social organisation of space. The first kind of lawfulness arises from mathematical constraints of possibility and links generative rules to morphological outcomes through emergent properties: consistent relational structures are the outcome of the repeated application of rules that in no way contain a prior description of these structures. The second kind of lawfulness bears upon the ways in which categorical distinctions and relations of control are mapped onto the objective properties of spatial patterns. The third kind of lawfulness concerns the generative effects of the organisation of space relative to social encounters and interactions.

Based on laws of the second kind, the invariant properties of the organisation of space contribute to the reproduction of social order. By contrast, laws of the third kind allow space to work generatively, that is, to support stable opportunities for co-presence and encounter which can lead to new social relationships of cooperation, communication or exchange. Thus, the dynamics of the production and reproduction of social structure, already addressed in *The social logic of space* (Hillier and Hanson 1984), are articulated in more precise terms. Lawfulness of the first kind may be seen to mediate between laws of the second and third kind. On the one hand, the generative rules applied to the production of built space may encode genotypical requirements associated with the distinction of categories or the imposition of control. On the other hand, however, the distribution of movement and the resulting opportunities for encounters are a function of the overall structure of space as it emerges from the interaction of generative rules and the objective laws of spatial possibility. Thus, social processes structure space, but the properties of the emergent structure of space respond to society and produce their own social effects.

Functionality passes through intelligibility

(Hillier, Burdett, Peponis and Penn 1987; Hillier 1996; Hillier and Iida 2005; Hillier and Vaughan 2007; Hillier 2012a)

The early work of Hillier and Leaman (1973, 1974) saw the relationship of human to artificial environments as primarily cognitive, with regard to not only the design and production but also the appropriation, inhabitation and use of built space. Furthermore, Hillier and Hanson (1984) assert that the question of how space is socially constructed is closely linked to the question of how spatial patterns are described. At the same time, and as a corollary of the emphasis on cognition, the earliest work as well as the latter (Hillier 1996) distinguish between intelligibility – the understanding of abstract organising principles within a built scheme – and visibility – the perception of spatial relationships.

But once space syntax analysis came to embrace the representation of perceptual affordances – such as the extension of uninterrupted lines of movement and sight, or the partitioning of open space into convex elements that associate co-presence with co-visibility – the question arose as to how the transition from the visible to the intelligible should be effectively conceptualised. In response, intelligibility was technically defined as the correlation between local properties of space, that can, by definition, more easily be perceived, and global properties that can only be understood through cognitive effort. The simplest and earliest technical definition of intelligibility was in terms of the correlation between the variables of connectivity and integration (closeness centrality) applied to the analysis of line representations of street networks, namely, axial maps. Hillier, Burdett, Peponis and Penn (1987) reported that the predictability of movement densities based on integration was stronger when the technically defined intelligibility of street networks (the correlation) was higher. Insofar as one of the functions of street networks is to distribute movement over an urban area, one could say that early research on natural movement demonstrated that functionality worked through intelligibility. The idea was taken up again as part of the formulation of the theory of natural movement (Hillier, Penn, Hanson, Grajewski and Xu 1993).

Hillier (1996) went on to argue that urban intelligibility, as technically defined, is more likely to occur when some of a settlement's lines of sight and movement are significantly longer than others. The longest lines of sight and movement usually intersect at wide angles to form continuous path sequences. This creates a cognitive ossature that facilitates navigation in urban areas. In other words, intelligibility is a function of the way in which spatial systems are scaled, including the scaling of perceptual affordances. The argument has inspired research in spatial cognition and wayfinding.[9]

For large systems, intelligibility cannot be reduced to the relationship between the local relations we perceive and the relations that we build in the mind as we come to understand spatial systems. A broader question is how our understanding of spatial structure encompasses different scales. Hillier (1996) addressed that question by considering the way in which urban areas, as conventionally identified and named, are related to the global structure of cities. He did this by looking at the relationship between integration at the scale of the city as a whole and integration within a restricted radius of analysis. Named local areas in Greater London such as Soho are characterised by a tighter correlation between the local radius of integration and global integration than prevails for the city as a whole or for the larger unit containing the smaller. So, for example, the City of London has a tighter correlation between radius integration and integration than Greater London, and the Leadenhall Market area has a tighter correlation than the City of London. Our collective recognition of distinct local areas is associated with objective morphological properties. The finding is significant because it suggests that the identifiability of urban areas is a function of the way in which local scales of connection are embedded within global scales. Identifiability can arise from intensifications of the relationship between local and global connectivity, and does not depend on disconnection or the imposition of boundaries.

In one of his later papers, Hillier (2012a) suggested that the city, as a human artefact, is structured so as to resonate with human cognition. The argument is in two main parts. First, the correlation between the distribution of movement and patterns of centrality based on turns or angular distance, rather than metric distance, suggests that cognitive ease, rather than minimisation of physical effort, is the driver of the correlation (Hillier and Iida 2005). Second, agent-based modelling experiments[10] show that movement is distributed more lawfully in systems where the longest lines of sight link up to form a quasi-continuous larger-scale network that weaves the local parts of cities into a coherent whole. There is a mutually reinforcing dialogue between the functional effects of the spatial structure of cities and the social production of the latter to resonate with the cognitive propensities of the human mind. The laws of space – the effects of the interface between the longer and shorter lines of sight and movement upon the distribution of closeness or betweenness centrality – mediate the link between cognition and natural movement.

Society as a spatio-temporal morphology of encounters situated in space

(Hillier and Hanson 1984; Hillier 1996; Hillier and Netto 2002)

The members of society are discrete individuals. Hillier proposes that one way to conceptualise society as a form realised in space and time is to consider it as a network of encounters and interactions periodically repeated over time. Encounters can arise in two ways. First, by virtue of propinquity, when people serendipitously meet each other as neighbours, as inhabitants of compact settlements or as users of shared open spaces between and around buildings. Second, by virtue of purposeful effort to overcome distance in order to maintain socially sanctioned relationships, such as kinship, or to participate in socially sanctioned activities, such as religious rituals. The members of most societies and social organisations participate in both kinds of encounter networks. The members of the faculty of universities, for example, may meet colleagues from their own or from other disciplines as they use common spaces on a university campus; and they may also meet colleagues from the same or allied disciplines by regularly attending regional, national or international conferences (Hillier and Penn 1991). Both kinds of meetings play an important role in their ability to develop ideas and to advance their respective fields; thus, both are essential for them to function as faculty members.

The recurring participation in meetings that require purposeful effort to overcome distance presupposes a continuing awareness of relevant identities (for example of one's position in a kinship system), of relevant beliefs (for example those attending a particular ritual) or of relevant organisational roles (for example one's position in a corporation that operates in different locations). In short, encounters across distance imply endorsement and assimilation of abstract descriptions of social membership (Hillier and Hanson 1984; Hillier and Netto 2002). This, in turn, may be reflected in the organisation of local space. For example, churches have certain common characteristics that help their congregations recognise themselves as members of a larger community across space. Thus, we come to expect that societies that are based on extensive networks of encounters across distance are likely to occupy local spatial syntaxes that comply with more extensive and complex organising principles.

The growth of large cities, and the concomitant potential for much denser and diversified patterns of encounter based on spatially based patterns of convergence, leads, according to Hillier, to the imposition of different kinds of organising principles. The issue brought forth by urban growth and increasing density is not so much the reproduction of abstract descriptions of social membership as the regulation and restriction of encounter probabilities. Constraints and transformations may be imposed on the design of the city itself, as, for example, when connected streets directly linked to building entrances are replaced by inward-looking enclaves. However, the need for regulation may also be expressed in the birth of new kinds of institutional buildings associated with political regulation and control and the positioning of such buildings within the urban fabric.

Hillier clearly recognised that societies are not homogeneous. Quite to the contrary, different social groups may be associated with distinct principles for reproducing themselves as encounter patterns in physical space. Some groups, usually the weakest in terms of social power or control, are sustained primarily by patterns of encounter based on propinquity or sustained co-presence thanks to patterns of local connectivity. Other groups, usually those that command greater resources, are associated with networks of encounter across space. These observations give rise to the idea that built space supports patterns of differential solidarity. The term 'differential solidarity' was coined by Hillier to characterise situations where an environment supports different modes of solidarity, such as Durkheim's organic and mechanical solidarity for different groups of people, thanks to the enmeshment of syntactic properties that can support each kind of solidarity without detriment to others.

One vivid expression of changes in mode of solidarity is provided by Hanson and Hillier's account of the transformation of terraced houses in London. Old working-class houses would often keep the door ajar, encouraging everyday visits by neighbours. Given restricted means, visits would tend to occur at times of the day other than those associated with eating meals. In the same houses, the parlour would be a visually protected best room set apart from everyday life and reserved for special occasions and infrequent visits from distant relatives. The same houses, when transformed under gentrification, would typically establish a *spectacular* visual connection between an open-plan interior and the street (for example, by keeping curtains open at night) while keeping the front door shut. Visits, by invitation only, would tend to be associated with socialisation over dinner.[11] Another vivid example, showing how the idea of differential solidarity can be applied to the study of cities, is Hillier's account of London (Hillier 1989). Buildings associated with economy and trade are located on main streets and major intersections. Churches associated with local communities take advantage of longer lines of sight throughout the fabric. Trade or craft guild headquarters are hidden in the interior of blocks to restrict access and exposure to those that have membership privileges. Thus, the morphology of the street network interacts with the patterns of localisation of different building types to accommodate different modes of solidarity associated with different principles of encounter.

If built space can indeed be organised to support the principles of encounter that are associated with different patterns of solidarity, an additional possibility arises. One group may have the power to impose spatial descriptions suitable for its own patterns of solidarity at the expense of descriptions that would allow the solidarity of other groups to flourish. This, argue Hillier and Hanson (1984), marks the transition from socio-spatial formations characterised by differential solidarity to formations that encompass socio-spatial class differences. For example, designs of urban space that restrict the potential encounters that would otherwise be possible based on street connectivity will privilege the modes of solidarity of the upper class at the expense of the modes of solidarity of the lower class. This is because the upper class commonly engages in patterns of socialisation founded on translocal networks of interest.

The emphasis on encounter patterns is intended to provide a spatially based view of society to complement the analysis of the social logic of built space. Thus, a natural bridge is created between space syntax and network-based approaches to the study of social phenomena.

Threads

The preceding sections have provisionally identified particular ideas, testable propositions and arguments in order to provide readers with signposts as they approach the work of Hillier. The way in which these ideas have continued to evolve and interact with each other, that is, the dynamics of Hillier's thought, are of course of much greater interest. Readers of this volume are likely to identify threads and themes that span over many years and bridge between conceptual development and empirically grounded research, critical thinking and philosophical argument.

One pervasive thread, for example, is the development of a theory of the city as a complex physical artefact that is constitutive of characteristic kinds of culture, economy and society. Here are the cornerstones of Hillier's approach. 1) The city is an agglomeration of buildings, serving different programmes of use, and accommodating different social groups, institutions or organisations. The particular object of interest, however, is not the agglomeration of, and interactions between, different specialised functions but rather the way in which they are connected through the street network, to form a composite spatial whole. The spatial structure of the city emerges over time, as local connections build up into global patterns of syntactic centrality. 2) The function of the spatial form of the city is to provide a long-term framework for interaction. Urban space provides an alternative basis for encounters as compared to formal social organisation. It has the potential to integrate across patterns of differentiation defined by social structure. 3) The study of urban spatial morphology leads to a distinction between scales of organisation. First, a primary foreground network of connected long lines of movement and sight which weaves the city together as a whole. Second, a background or infill network comprising the shorter street lines that make up the patchwork of local urban areas. The distinction and interface between different scales is fundamental to the spatial structure of the city and a universal morphological regularity. 4) The foreground network of long lines supports patterns of economy and exchanges between products, information or specialised bodies of knowledge and social practice. The background network is structured to support patterns of community subject to the prevailing social norms that characterise individual cultures. Cities are thus characterised by differential forms of solidarity. They accommodate encounter patterns subject to different organising principles. Some take advantage of relations of propinquity and connection; others result from social norms, formal patterns of organisation or institutionalised practices. Sometimes, one social group has the power to impose forms of spatial organisation that do not support the patterns of co-presence and encounter that characterise the social solidarity of another group. This gives rise to fundamental spatial inequalities. 5) We can identify different types of urban structure by the sub-shapes that emerge according to the distribution of the most integrated lines at various scales. In cities or parts of cities associated with vivid patterns of enterprise and exchange, the network of the most integrated streets takes the shape of a 'deformed wheel' that links the middle to the periphery and all the parts to each other. Cities or parts of cities associated with restrictions over patterns of interaction and exchange tend to be characterised by tree-like hierarchical patterns of integration, or by disconnected local centres. 6) Patterns of movement and potential encounter are distributed according to syntactic centrality, thus creating interfaces between quieter and busier areas, between local interactions and the awareness of global flows and connections. The distribution of movement indexes the power of urban spatial configuration to produce its own structuring effects upon social co-presence and interaction; thus, to potentially generate new patterns of coordination, collaboration and exchange. 7) The distribution of movement becomes a dimension of urban economy. Land uses such as retail, which benefit from the 'passing trade', tend to be located on better-connected streets, thus multiplying the effect of purely spatial centrality. Commercial attraction becomes distributed not only according to distance from the central business district or other similar hubs but also according to the syntax of the primary street network. 8) The association between the distribution of movement and the pattern of syntactic centrality is founded on the resonance between city form and spatial cognition. When distances are measured by path turns or by aggregate path rotation, local connectivity is correlated with local centrality and in turn local centrality is systematically interfaced with global centrality. The predictability of the distribution of movement is greater when these correlations are stronger. The properties revealed by the correlations, however, are not mathematically necessary for all possible street networks but emerge only thanks to the way in which the longest lines of movement and sight link up to create a primary network at various scales.

The cornerstones of Hillier's approach to the city define a multidimensional conceptual space. A reader can trace lines that link the harnessing and organisation of the perceptual affordances of space to the production of the emergent structural properties of urban layouts; the functional effects of spatial structure with respect to the co-presence and encounter to the encoding of the abstract principles that govern social interaction and modes of solidarity; and economic performance to cognitive legibility.

There are many other threads that a reader might choose to unravel. The development of analytical concepts, techniques of computational analysis and measures of spatial organisation is one of them. The foundations of space syntax analysis are presented in Hillier and Hanson's *The social logic of space* (1984), and Hillier's *Space is the machine* (1996). There are, however, additional key papers of which Hillier is not a primary author, or co-author, including those by Alasdair Turner, Nick Sheep Dalton, John Peponis and Michael Batty. In this context we wish to recognise a number of people associated with the development of software that enabled not only the consistent application of space syntax analysis but also Hillier's 'thought experiments' linking the development of formal concepts to the study of morphological possibility: Paul Coates, who developed the first simulations of beady ring processes of growth, and the first software for the space syntax analysis with graphic output in the 1970s and 1980s; Nick Sheep Dalton, who developed *Axman*, the first easy-to-use software with graphic input and output, in the 1980s and 1990s; and Alasdair Turner, whose contributions include the development of the space syntax software most commonly used around the world, *DepthMap;* the integration of visibility analysis into space syntax; and the introduction of agent simulations of movement. Readers may trace multiple lines of thought that stem from these developments in the discipline's research software. For example, from topological concepts of relation and connection to representations of spatial relationships that are sensitive to projective and Euclidean geometry; also, from pure graph-based measures that describe the relationships of dimensionless elements, to graph-based measures that incorporate weightings by length and angle. They may also trace lines of development proceeding from the simple invention or adoption of measures to the systematic exploration of the relationship-subtending properties of spatial configurations, as, for example: first, the interdependence of the measure of normalised mean depth (closeness centrality) to the number of elements comprised by a system, a finding that gave rise to the size-adjusted measure of closeness centrality known as *integration* (Hillier and Hanson 1984); second, the interrelation between the distribution of line length and correlation between connectivity and integration in line map representations of street networks, a finding central to the theory that urban form is constructed so as to support intelligibility; third, the covariation of closeness and betweenness centrality offered by Hillier, Yang and Turner (2012), which gave rise to the composite measure known as *relativised choice*.

One thread of particular interest that remains to be unravelled concerns the relationship between the theories of the social logic of space and architectural design. The understanding that space syntax analysis can be applied to clarify design aims that bear upon social function, and to test and develop design proposals in the light of such aims, does not exhaust the matter. Particularly in the case of urban design, it is clear that space syntax has and can continue to be so applied (Hillier, Hanson, Peponis, Burdett and Hudson 1983; Hillier 1993), especially where the projects, and the patterns of movement that that they are likely to engender, are to be integrated into their urban context for the benefit of their surroundings as well as of the projects themselves. Should we, therefore, consider space syntax as a particular kind of expertise that can be brought to bear to strengthen design proposals with respect to significant social functions? If so, space syntax's contribution to architectural practice is undoubtedly of value. But how does such a specialised body of knowledge resonate with the broader perspectives regarding design pre-structures, the relation of knowledge and design, and architecture as the introduction of theory to building, perspectives that underpin and continue to be developed in Hillier's work over the years?

Hillier (1993) himself offers compelling answers to some of these questions. For example, when applied to precedents, syntactic analysis leads us to abstract principles of urban spatial organisation and to identify significant structural properties. As a consequence, our knowledge of what is possible in urban design becomes independent of the individual characteristics of any particular precedent. In this sense, we are freed from the constraints of repetition and enabled to imaginatively endorse and interpret abstract structural relationships. Space syntax fulfils the aims endorsed in 'Knowledge and design' (Hillier, Musgrove and O'Sullivan 1972), to strengthen both the knowledge base of design and design intuition. It articulates theoretically explicit and testable 'design pre-structures' such as those that make design possible according to Hillier and Leaman (1974).

But Hillier also recognises a much more subtle question. Design addresses multiple requirements that are distinct in nature, such as the requirements associated with environmental performance and the requirements associated with behavioural and cognitive performance. These requirements, however, interact in any particular design, as the satisfaction of each has to become congruent with the satisfaction of others (Hillier 1993). Thus, while the theories that address the functions of buildings lead to the identification and understanding of distinct abstract principles, design practice leads to concretion, the congruent application of these principles in a particular case. The question arises: what, if anything, can space syntax help us learn regarding the consequences of such necessary concretion? Given the interplay between the analysis of structural properties and the analysis of perceptual affordances, identified in an earlier section of this introduction, Hillier makes a crucial observation, taking house designs as a case in point. The structural properties associated with the satisfaction of cultural exigencies can be realised in designs that also provide distinct perceptual affordances. In turn, the perceptual affordances can be tuned to the satisfaction of aesthetic aims and aesthetic functions. If in turn we think of the satisfaction of cultural requirements as a key end in design, and of the shaping of perceptual affordances as the means, then architects can, to use Hillier's words (1993), 'modify the ends by re-expressing them as part of a richer cultural realm'. In the designs of Le Corbusier and Adolf Loos, 'the cultural and functional differentiation of space is the social meaning and the spatial means is the spatial aesthetic'. Loos organises visual fields to enhance the connections of one behavioural setting to others. Le Corbusier organises visual fields to generate dramatically changing views over movement trajectories. To use the metaphor of Dickon Irwin, at the time an MSc student, whose work Hillier references, Le Corbusier creates cinematic space, a montage of sculpturally rich views, while Loos creates theatrical space, a space where a participant in one setting is also a spectator relative to the participants in others.

Thus, syntactic analysis can be applied towards understanding the significant consequences of the interaction of spatial desiderata. As a result, it can lead to a comparative understanding of modalities of concretion. This particular thread of argument, initiated by Hillier, has not been vigorously pursued in his subsequent work. Yet it exemplifies the way in which the pursuit of connections between the ideas presented in different papers can lead to exciting extensions of space syntax into core concerns of culturally but also theoretically informed architectural criticism. One can see, in such potential extension, the fulfilment of the challenge posed in 'Quite unlike the pleasures of scratching' (Hillier 1985): to link the aesthetic and the social meaningfulness of architectural form. We hope that the selection of papers included in this volume will encourage and facilitate readers to identify the threads of ideas that relate to their own interests.

Hillier papers cited

Hillier, B. 1985. 'The nature of the artificial: the contingent and the necessary in spatial form in architecture'. *Geoforum*, 16, 163–178.
Hillier, B. 1985. 'Quite unlike the pleasures of scratching: Theory and meaning in architectural form'. *9H*, 66–72.
Hillier, B. 1988. 'Against enclosure'. In: Teymur, N., Markus, T. A. and Wooley, T. (eds.), *Rehumanising housing.* Oxford: Butterworth.
Hillier, B. 1989. 'The architecture of the urban object'. *Ekistics*, 56, 5–21.
Hillier, B. 1993. 'Specifically architectural theory: A partial account of the ascent from building as cultural transmission to building as theoretical concretion'. *Harvard Architecture Review*, 9, 8–27.
Hillier, B. 1996. *Space is the machine*. Cambridge: Cambridge University Press.
Hillier, B. 1997. 'Cities as movement economies'. In: Droege, P. (ed.), *Intelligent environments: Spatial aspects of the information revolution, North Holland, Amsterdam, The Netherlands.* Amsterdam: Elsevier.
Hillier, B. 2002. 'A theory of the city as object: Or, how spatial laws mediate the social construction of urban space'. *Urban Design International*, 7, 153–179.
Hillier, B. 2012a. 'Studying cities to learn about minds: Some possible implications of space syntax for spatial cognition'. *Environment and Planning B: Planning and Design*, 39, 12–32.
Hillier, B. 2012b. 'The genetic code for cities: Is it simpler than we think?' In: Portugali, J., Meyer, H., Stolk, E. and Tan, E. (eds.), *Complexity theories of cities have come of age: An overview with implications to urban planning and design.* Berlin and Heidelberg: Springer, 129–152.
Hillier, B. 2014. 'The generic city and its origins'. *Architectural Design*, 36, 100–105.
Hillier, B. 2016. 'The fourth sustainability, creativity: Statistical associations and credible mechanisms'. In: Portugali, J. and Stolk, E. (eds), *Complexity, cognition, urban planning and design*. Berlin and Heidelberg: Springer, 75–92.
Hillier, B., Burdett, R., Peponis, J. and Penn, A. 1987. 'Creating life: Or, does architecture determine anything?' *Architecture et Comportement/Architecture and Behaviour*, 3, 233–250.
Hillier, B. and Hanson, J. 1984. *The social logic of space*. Cambridge: Cambridge University Press.
Hillier, B., Hanson, J. and Graham, H. 1987. 'Ideas are in things: An application of the space syntax method to discovering house genotypes'. *Environment and Planning B: Planning and Design*, 14, 363–385.

Hillier, B., Hanson, J. and Peponis, J. 1984. 'What do we mean by building function?' In: Powell, J. A., Cooper, I. and Lera, S. (eds.), *Designing for building utilization*. London: Spon.
Hillier, B., Hanson, J., Peponis, J., Hudson, J. and Burdett, R. 1983. 'Space syntax: A different urban perspective'. *The Architects' Journal*, 178(48), 47–63.
Hillier, B. and Iida, S. 2005. 'Network and psychological effects in urban movement'. In: Cohn, A. and Mark, D. M. (eds.), *Spatial information theory*, *Lecture notes in computer science*, no. 3693. Berlin and Heidelberg: Springer, 475–490.
Hillier, B. and Leaman, A. 1972. 'A new approach to architectural research'. *RIBA Journal*, 79, 512–521.
Hillier, B. and Leaman, A. 1973. 'The man–environment paradigm and its paradoxes'. *Architectural Design*, 78, 507–511.
Hillier, B. and Leaman, A. 1974. 'How is design possible?' *Journal of Architectural Research*, 3, 4–11.
Hillier, B. and Leaman, A. 1975. 'The architecture of architecture'. In: Hawkes, D. (ed.), *Models and systems in architecture and building*. London: The Construction Press.
Hillier, B. and Leaman, A. 1976. 'Architecture as a discipline'. *Journal of Architectural Research*, 5, 28–32.
Hillier, B., Leaman, A., Stansall, P. and Bedford, M. 1976. 'Space syntax'. *Environment and Planning B: Planning and Design*, 3, 147–185.
Hillier, B., Musgrove, J. and O'Sullivan, P. 1972. 'Knowledge and design'. In: Mitchell, W. J. (ed.), *Environmental design: Research and practice*. Los Angeles: The University of California Press, 29.3.1–29.3.14.
Hillier, B. and Netto, V. 2002. 'Society seen through the prism of space: Outline of a theory of society and space'. *URBAN DESIGN International*, 7, 181–203.
Hillier, B. and Penn, A. 1991. 'Visible colleges: Structure and randomness in the place of discovery'. *Science in Context*, 4, 23–50.
Hillier, B., Penn, A., Hanson, J., Grajewski, T. and Xu, J. 1993. 'Natural movement: Or, configuration and attraction in urban pedestrian movement'. *Environment and Planning B: Planning and Design*, 20, 29–66.
Hillier, B. and Tzortzi, K. 2006. 'Space syntax: The language of museum space'. In: Macdonald, S. (ed.), *A companion to museum studies*. Oxford: Blackwell.
Hillier, B. and Vaughan, L. 2007. 'The city as one thing'. In: Vaughan, L. (ed.), 'The syntax of segregation'. Special issue, *Progress in Planning*, 67, 205–230.

Notes

1. J. Piaget and B. Inhelder, *The child's conception of space* (London: Routledge and Kegan Paul, 1956).
2. L. March and P. Steadman, *The geometry of environment: An introduction to spatial organisation in design* (London: RIBA Publications, 1971).
3. J. Peponis, J. Wineman, S. Bafna, M. Rashid and S. H. Kim, 'On the generation of linear representations of spatial configuration', *Environment and Planning B: Planning and Design* 25 (1998): 559–576.
4. A. Turner, A. Penn and B. Hillier, 'An algorithmic definition of the axial map', *Environment and Planning B: Planning and Design* 32 (2005): 425–444.
5. M. L. Benedikt, 'To take hold of space: Isovists and isovist fields', *Environment and Planning B: Planning and Design* 6 (1979): 47–65.
6. A. Turner, M. Doxa, D. O'Sullivan and A. Penn, 'From isovists to visibility graphs: A methodology for the analysis of architectural space', *Environment and Planning B: Planning and Design* 28 (2001): 103–121.
7. N. Dalton, 'Fractional configurational analysis and a solution to the Manhattan problem', in *Third International Space Syntax Symposium*, edited by J. Peponis, J. Wineman and S. Bafna (Atlanta, GA: Georgia Institute of Technology, 2001), pp. 26.1–26.13; A. Turner, 'Angular analysis', in *Third International Space Syntax Symposium*, edited by J. Peponis, J. Wineman and S. Bafna (Atlanta, GA: Georgia Institute of Technology), pp. 30.1–30.11.
8. For example, J. Peponis, 'The spatial culture of factories', *Human relations* [New York] 38 (1985): 357–390.
9. See, for example, K. R. Barton, D. Valtchanov and C. Ellard, 'Seeing beyond your visual field: The influence of spatial topology and visual field on navigation performance', *Environment and Behavior* 46 (2014): 507–529.
10. Hillier references Turner, 'Angular analysis', but the reader may also want to consult A. Turner, 'DepthMap: A program to perform visibility graph analysis', in *Third International Space Syntax Symposium*, edited by J. Peponis, J. Wineman and S. Bafna (Atlanta, GA: Georgia Institute of Technology), pp. 31.1–31.9; and A. Turner and A. Penn, 'Encoding natural movement as an agent-based system: An investigation into human pedestrian behaviour in the built environment', *Environment and Planning B: Planning and Design* 29 (2002): 473–490.
11. J. Hanson and B. Hillier, 'Domestic space organization: Two contemporary space-codes compared', *Architecture et Comportement/Architecture and Behaviour* 2 (1982): 5–25.

1 Knowledge and design (1972)

Introduction to 'Knowledge and design'

Ruth Conroy Dalton

This paper should have been far more influential than it was, and I will return to this statement at the end of this introduction, suggesting reasons as to why I think this might have been the case. The paper is about the surprisingly intricate and complex relationship between architectural knowledge, research and design. The term 'knowledge', here, refers to the holistic understanding that informs design and to which research of the *right kind* contributes; it implies a synthesis of information, experience and insights.

The introduction discusses the historical development of ideas in architectural research and demonstrates why they existed at the time of writing (and arguably still exist to some extent today), and hence provides evidence for an 'applicability gap' between architectural research undertaken and designers' needs for information. It describes the problems inherent in architectural research, as it had been previously conceived, demonstrating that architectural research had, hitherto, not been fit for purpose because of its emphasis on making design a systematic, problem-solving activity. Such a mismatch between the research *undertaken* and the research *needed* (by designers) is attributed to a fundamental misunderstanding of the process of design.

The paper then examines the design process to demonstrate *how* the activity of design has been misunderstood and, as a direct corollary of this, how the focus of architects' education, as conceived at the time, was misplaced. This lack of understanding of the design process meant that: i) we do not know *what* knowledge designers need; ii) we do not know *when* to apply that knowledge; iii) we do not know *how* to apply that knowledge and, therefore, iv) we do not know *how to teach* design.

This problem articulation leads to the next section linking the philosophical foundations and paradigmatic shifts in architectural research to the creation and evolution of scientific (not architectural) knowledge. The authors assert that science, and by extension scientific architectural research, is not about producing independent, factual knowledge. Scientific endeavour is a primarily cognitive activity that aims to provide a working understanding of the world and, in this sense, has parallels with design activity, which aims to build the world whilst anticipating the consequences of any design decisions. The distinctive characteristic of science is that the propositions through which we understand the world are empirically testable and refutable (and by extension, this should also be true of design).

Next, the authors characterise past paradigms of design activity as assuming, wrongfully, that design was a process of analysis-synthesis. The process of analysis-synthesis can be considered 'inductive' as it involves reasoning from specific observations or examples to form general conclusions. In the context of analysis-synthesis, the process typically starts with an analysis of individual components, data or observations to identify patterns or commonalities leading to the synthesis of a general understanding or solution – it is both logical and accretive and, as such, 'information . . . generate[s] the solution'. However, as the authors point out, this is an entirely fallacious model regardless of how seductive it seems.

They then discuss the role that reflexivity plays in the process of pre-structuring design problems. The authors introduce the idea of an 'instrumental set' which, essentially, consists of resources that can be brought to bear on a design problem and includes any tools and the materials to which the tools are applied. However, and this is something I find particularly intriguing, they regard the designer's cognition as part of this 'instrumental set'. The familiar solution 'types' that drive the designer's approach to the problem are *as much of a tool* as a simple pencil or chisel. So are the 'codes', or the assumptions, principles and symbolic patterns that translate abstract design desiderata into solution types. They point out that this instrumental set can be applied *in exactly the same way* to architectural research as to the act of design itself.

Next, they consider the interplay of research and knowledge types. They categorise different types of research to consider how they map onto different elements in the 'instrumental set' and/or the 'designer's field', what outcomes or deliverables each can produce and what utility they may ultimately have for the designer. They conclude that it is the research that seeks to make explicit, test and enrich the codes (by which designers link design requirements to possible formal responses) that are the most useful of the four.

This leads to the important idea of conjecture-driven design. Drawing parallels with scientific methodology, they emphasise the vital importance of conjecture in the design process. They argue that conjecturing, alongside rigorous testing, is fundamental to the progress of *both* science and design. The designer's cognitive capability (and, importantly, self-awareness of this), including knowledge of instrumental sets, solution types and design codes, plays a crucial role in this conjecture.

They conclude by exploring this newly proposed definition of architectural research. As all buildings can be treated as activity/behaviour-, climate-, resource- and cultural modifiers (a definition explored more extensively in other papers in this volume), research can be broadly categorised according to which aspect of building function is addressed. The authors show how the impact of broader scientific fields is only realisable through integrative theories that progressively shape the response to these four fundamental functions of buildings, hence treating the contributory disciplines as extensions of the cognitive foundation of design.

Since reading this paper again, I have been wondering why it never became as well-known and influential as other, later academic texts presenting similar ideas. I think the text poses a formidable challenge owing to its intricate terminology; it employs specialised language and concepts that can render it challenging to follow the subtleties of the authors' arguments. Furthermore, the text covers a range of complex ideas regarding the essence of design, the cognitive processes of designers and the interplay between theory and practice, each of which can be challenging. But primarily, it is the text's *density* of ideas that contributes to its demanding nature. It encapsulates a multitude of layered concepts, making it a formidable task for readers to navigate. This, I believe, is a key reason why the text has never enjoyed widespread recognition. However, its very richness, while presenting a challenge, also renders it exceptionally rewarding for those willing to tackle it.

Knowledge and design

Bill Hillier (and Adrian Leaman), with John Musgrove and Pat O'Sullivan

Originally published: Hillier, B. (and Leaman, A.), with Musgrove, J. and O'-Sullivan, P. 1972. 'Knowledge and design'. In: Mitchell, W. (ed.), *AIA-Architect-Researchers' Conference. Environmental Design: Research and Practice. Proceedings of the Third Conference of the Environmental Design Research Association*. Los Angeles: University of California Press, 245–264.[1]

Introduction

The desired outcome of environmental research is usually seen in terms of scientific knowledge and rules which decrease the designer's reliance on intuition and rules of thumb in solving problems. The corollary of this idea is that design methods must be systematised in order to assimilate such knowledge. These two notions form the basis of a kind of paradigm for environmental research and its usefulness. It is argued that this paradigm is based on simplistic notions about knowledge and about design, that design problems are essentially pre-structured[2] both by constraints and by the designer's own cognitive map, and that solutions are only intelligible if this is fully recognised. Design proceeds by conjecture-analysis rather than by analysis-synthesis. It is argued that if research is to make an impact on design, it must influence designers at the pre-structuring and conjectural stages. The idea that research should produce knowledge in the form of packaged information, coupled to rationalised design procedures, is therefore inadequate. The aim of research should be seen more in terms of providing designers with a stronger theoretical, operational and heuristic basis from which to conjecture, rather than in terms of knowledge to determine outcomes.

This paper is about knowledge and design. Before you reach for your hats, we should add that it is intended to be a technical not a theological contribution to the subject. We offer it, not because we have no research to report, or because we think this conference is in need of philosophic homilies, but because we believe we are up against some fundamental limitations in the arguments which have been used to justify research and to define its tasks.[3,4]

Research of one kind or another has now a longish history in building. By and large, this increased investment in research has proceeded side by side with a marked deterioration in the quality of building. A serious 'applicability gap' appears to exist. Regardless of the quality of research work itself, the history of attempts to link research to improvements in environmental action is largely one of confusion and failure.[5]

When the concepts of 'environmental research', as opposed to technologically oriented building research, began to emerge a decade or so ago, there seemed to be good reasons for hoping that these new concepts in research would lead to new relationships between research and action. At that stage, when the ground was being cleared for the expansion of environmental research, programmatic statements took a clear line.

Ten years ago, when the ground was being cleared for a great expansion of architectural research activity, programmatic statements took a clear line. Design was a problem-solving activity, involving quantifiable and non-quantifiable factors. Research, it was thought, should bring as many factors as possible within the domain of the quantifiable, and progressively replace intuition and rules of thumb with knowledge and methods of measurement. This process would never be complete. Non-quantifiable elements would remain. In order to assimilate such knowledge and use such tools as we were able to bring to bear on design, the procedures of designers would have to be made more systematic. Because the education of architects was broad and shallow, and because they were concerned with action rather than knowledge, they could not be expected to generate new knowledge for themselves. This was the job of 'related' disciplines, whose concern was the advancement of knowledge. Architects, on the other hand, knew about design, and should make systematic design their research focus. Otherwise, their contribution to research lay in technological development, or as members of multi-disciplinary teams, in defining the problems for others to solve.

The educational consequences of these notions were that schools of architecture and planning were to be located in an educational milieu containing a rich variety of related disciplines, and students were to be well grounded in each of them. The core of the architectural course would still be design and, at the academic level, this meant increased concentration upon systematic methods. Students would be taught to analyse problems, and to synthesise solutions.

A few voices crying in the wilderness that architecture contained its own fundamental disciplines could not stop the onward march of these simple and powerful ideas, and by and large, they still hold the stage today. But if these are to be the paradigmatic ideas by which we define our subjects and link them to action, then today's landscape (although promising in that other disciplines are developing their latent 'environmental' interests) must appear depressing. Systematic design studies are in disarray. Increasing numbers of research workers, including architects, are moving into the areas previously called 'unquantifiable'. There is a widespread feeling that an 'applicability gap' has developed between research and design. Design is still led by the nose by technology, economics and imagistic fashions. The human sciences and architecture are still at loggerheads. Education, with few exceptions, has not managed to develop a radically new capability in the problem-solving power which students bring to design.

In fact, we are far from pessimistic about the progress of architectural research, largely because a great deal is now happening that cannot be explained in terms of the ideas we have outlined above. The situation has outstripped the paradigm which gave birth to it. But the intelligibility of the situation is poor, perhaps because it is inconsistent with the paradigm. We require some radical overhaul of its assumptions – particularly those to do with the relationships between knowledge and design, and the presupposed polarities (for example, rationalism/intuitionism) – along with a new effort to externalise the dynamics of the new situation. To us this seems to be an essential step before the 'applicability gap' is compounded by a 'credibility gap' arising from the gulf between what is expected of research, and what research appears to be offering.

Perhaps the simplest way of introducing what we have to say is by drawing an analogy with the slow but decisive shift in philosophy and scientific epistemology over the past half-century or so. Implicit in both the rationalist and empiricist lines of thought was the notion that in order to get at truth, preconceptions must be eliminated or at least reduced to the minimum. Rationalism began its long history by proposing *a priori* axioms whose truth was supposed to be self-evident; empiricism relied on the neutrality of observation. Since the early part of this century, developments in such areas as psychology, meta-mathematics, logic and the philosophy of science have combined to show that both of these are impossible and unnecessary to an account of scientific progress. Far from being removed from the field of science, the cognitive schema by which we interpret the world and pre-structure our observations are increasingly seen to be the essential subject matter of science. The question is not whether the world is pre-structured, but how it is pre-structured.

Too often these developments appear to have escaped the attention of scientists working in the environmental field and designers interested in research and looking to research for solutions to problems. This is particularly unfortunate because the idea of pre-structuring has immediate and fundamental applications in design. We cannot escape from the fact that designers must, and do, pre-structure their problems in order to solve them, although it appears to have been an article of faith among writers on design method (with a few exceptions[6]) that this was undesirable because unscientific. The nub of our argument is that research in the field of the built environment and its action systems should see as its eventual outcome, and point of aim the restructuring of the cognitive schemes which designers bring to bear on their tasks, not in terms of supplying 'knowledge' as packaged information to fit into rationalised design procedures, but in terms of redefining what those tasks are like, and using the heuristic capability of scientific procedures to explore the possible through a study of the actual. It is our view that the notion that well-packaged knowledge coupled with a logic of design can lead to radically better artefacts, on the evidence we have, should be relegated to the realm of mythology. But in arguing our case we would like to say a little more about why we think modern scientific epistemology has an important bearing on design and meta-design (which we will argue is probably the simplest and most adequate characterisation of design research) and why it can help us reconstitute our paradigmatic notions about the subject.

Fifty years ago, it was still possible to think of science as ultimately constituting a set of signs which, at the most rudimentary level, would bear a one-to-one correspondence with atomic facts, and that these could eventually be combined by the laws of induction and verification into a pyramid of laws of greater and greater generality. Scientists, on the whole, believed this to be the case, and philosophers concerned themselves to show how it would be accomplished.

The overthrow of the Newtonian account of the universe, previously taken as the paradigm of positive knowledge, arrived at by observation and induction (as described by Newton himself), threw scientific epistemology into a crisis, the effects of which are still with us. We will give a brief account of this later on. Shortly afterwards, even more remarkable and undermining developments took place in the foundations of mathematics and logic. Gödel showed, by his incompleteness theorem, that 'the construction of a demonstrably consistent relatively rich theory requires not simply an "analysis" of its "presuppositions", but the construction of the next higher theory with the "effect"', to continue quoting Piaget, that

> 'Previously it had been possible to view theories as layers of a pyramid, each resting on the one below, the theory at ground level being the most secure because constituted by the simplest means, and the whole poised on a self-sufficient base. Now, however, "simplicity" becomes a sign of weakness and the fastening of any story on the edifice of human knowledge calls for the construction of the next higher theory. To revert to our earlier image, the pyramid of knowledge no longer rests on its foundations but hangs by its vertex, an ideal point never reached and, more curious, constantly rising.'[7,8]

This is of vital importance, not simply because it demonstrates the inherent limitations of formalism, and the impossibility of such notions as the class of all classes, or the single unified science, but because it demonstrates that there is a necessary hierarchy which limits what we can mean by knowledge – the hierarchy of meta-theories and meta-languages, independent of (we can think of it as orthogonal to) the hierarchy of levels of integration of phenomena in the 'real' world, which constitutes the formal basis of most scientific disciplines. Any cognitive formalisation takes a lower order formalism for its object and can itself become the object of a higher formalism. To quote Piaget again:

> 'The limits of formalism can, more simply, be understood as due to the fact that there is no "form as such" or "content as such", that each element – from sensory motor acts through operations to theories – is always, simultaneously form to the content it subsumes and content for some higher form.'[9]

If we accept that the idea of a monumental edifice of knowledge, descriptive of the world in its account of facts and explanatory of them in terms of theories of increasing generality, has to be given up, what have we left? Have we not effectively debunked the idea of knowledge? Having got rid of positivism, are we left with pure relativism? Intuitively, we feel that such a retreat cannot account for the success of science in improving our understanding of the world and our capacity to modify it. If we adopt a position of pure, philosophic relativism, then relativity (in theoretical physics – we are short of terms here) and the atomic bomb appears as a kind of epistemological paradox. If, on the other hand, we accept that there are strong reasons for rejecting both positivism and pure philosophic relativism, then where do we go? It seems that, as with Scylla and Charybdis, we cannot escape the one without falling into the other.

It is against this background that the achievement of scientific philosophers like Karl Popper, Thomas Kuhn and Imre Lakatos take on their full stature. Popper has demonstrated that a logic of induction and the principle of verification, previously the twin pillars of positivist science, were both unattainable and unnecessary, and that science could be contained within a hypothetico-deductive scheme:[10] Kuhn suggests a changing epistemological paradigm, within which science can operate as a puzzle-solving activity until the next revolutionary 'paradigm switch'.[11] Lakatos reconstructs science as conflicting sets of interrelated theories (on a smaller scale and more volatile than Kuhn's paradigms), retaining the idea of a 'negative heuristic theoretical core' and a 'positive heuristic' puzzle-solving area, each of which exhibits, at any time, either a 'progressing' or a 'degenerating' problem shift according to whether or not it is able to predict new phenomena within its basic theories without having to add *ad hoc* hypotheses to account for newly discovered phenomena.[12] Then we have a reconstruction of science which is able, in a highly nonlinear way, to account for its own continuity, as well as some rational justification for using the word 'knowledge' perhaps, to use Popper's expression, as 'piles in the swamp', the swamp being essentially the infinite regress of meta-theories and meta-languages.

The simplest reconciliation of these lines of thought in meta-mathematics and the philosophy of science is to state frankly that the object of science is cognition, and that it is the stratagems of science that are directed towards the real or empirical world. More precisely, we could say that science is about 'remaking cognition', it being clear that if we were satisfied with our cognitive codes for deciphering the world, we would not have science. This seems to us an adequate resolution of the old philosophical problem of whether the 'world out there' or our perception of it is the more real. Such a definition is implicit in the work of psychologists, like Kelly, who characterise everyday behaviour by analogy with scientific behaviour.[13] It is a small step to reverse the argument, and it allows us to account not only for the preoccupation of science with the empirical, but also for the fact that some advanced areas of science – notably certain branches of theoretical physics – have had no means of contacting the world for about 40 years. We would hardly be satisfied with a characterisation of science which relegated theoretical physics to the realm of metaphysics.

How does all this help us with architectural research? First, it should be clear that once we move away from the establishment of basic criteria set up with a view to avoiding physical discomfort (which we knew how to do anyway in pre-scientific days) then we can avoid a lot of misconception about the status of 'knowledge' in design. Second, we can begin to see the problems raised by the paradigm for research in architecture that we outlined at the beginning of the paper. Third, it provides us with a better method of making fertile analogies between, and thus in connecting the activities of, scientists and designers.

The paradigm we suggest as underlying most current research activity in architecture appears to be based on two notions about science that take no account of the developments that we have outlined: the notion that science can produce factual knowledge, which is superior to and independent of theory; and the notion of a logic of induction, by which theories may be derived logically from an analysis of facts. In the paradigm, these two notions appear to constitute the fundamental assumptions on which the whole set of ideas is founded: first, that the role of scientific work is to provide factual information that can be assimilated into design; second that a rationalised design process, able to assimilate such information, would characteristically and necessarily proceed by decomposing a problem into its elements, adding an information content to each element drawn, as far as possible, from scientific work, and 'synthesising' (that is, inducting) a solution by means of a set of logical or procedural rules.

So far, we have suggested very theoretical reasons why such ideas would not be viable or realisable. But equally, from the more practical point of view of the designer or the student, the ideas – or more precisely the operational consequences that flow directly from them – appear even more unviable. Designers are left to make their own links with research by assimilating 'results' and quantification rules, and to evaluate them as they appear without guidance on priorities or patterns of application. The designer's field thus becomes *more* complex and *less* structured. It follows that if a designer cannot make use of this 'information', he is forced to the conclusion that it is because his procedures are not systematic enough, with the result that if he tries to improve himself, he immediately becomes preoccupied with means at the expense of ends.

Similar consequences flowed from these twin paradigmatic assumptions in architectural research itself. For example, building science as a university discipline tended to remain separate and independent of the design disciplines, usually as a research-oriented, service-teaching department, sometimes even generating the packages of knowledge that were to fit into the rationalised design procedures. In trying to formalise the process, designers were forced into developing concepts like 'fit' and 'optimisation' simply in order to complete the line of logic by which 'synthesis' could be accomplished, even though such notions are highly artificial in terms of what buildings are really like and are actually refuted by considering buildings as time-dependent systems rather than as once-and-for-all products.

Our negative aim in this paper was to try to show why the advance of research related to design has so far appeared to progress in parallel with deterioration in the acceptability of the designed product – and this, in the UK, in spite of two decades of excellent work by such bodies as the Building Research Station[14] and government departments, well disseminated in intelligible form and often containing mandatory requirements. We hope that we have shown that there are both theoretical and practical reasons why such a state of affairs should not surprise us. If the present paradigm is unworkable in its essentials, what can we put in its place? We have to preface our proposals with some suggestions about the nature (the actual nature as well as the desirable nature) of design activity.

Part 1

It is not hard to see why the analysis-synthesis, or inductive, notion of design was popular with theorisers and even with designers as a rationalisation of their own activities. The architectural version of the liberal-rational tradition was that designs should be derived from an analysis of the requirements of the users, rather than from the designer's preconceptions. It is directly analogous to the popularity of induction with scientists who were anxious to distinguish their theories as being derived from a meticulous examination of the facts in the real world. The point we are making in both cases is not that the ideas are immoral or fundamentally deceptive – scientists do describe meticulously the 'facts' of the situation, and designers do pay attention to the details of user needs – it is that they are theoretically untenable and unnecessary, and as a result, practically confusing.

The first point we would like to make about our version of science in relation to design is that if scientists really operate by a kind of dialectic between their pre-structuring of the world and the world as it shows itself to be when examined in these terms, then why should such a procedure be thought unscientific in design? Why not accept that only by pre-structuring any problem, either explicitly or implicitly, can we make it tractable to rational analysis or empirical investigation?

The second point is also in the form of a question. If rationality in design is not to be characterised in terms of a procedure that allows the information to generate the solution, then in what terms can it be characterised? Is it a redundant notion? Is there any alternative to the mixture of intuitive, imitative and quasi-scientific procedures which appear to characterise design as it is carried out? We would like to work towards answers to both of these questions by using some of the ideas we have discussed in a kind of thought experiment about the nature of design.

First, source observations about reflexivity (cognitive activity making itself its own object; or part of its object) and meta-languages and meta-theories (cognitive activity making other cognitive activity its object). These, it would appear, have clear parallels at the social level, in terms of the progressive differentiation of roles, especially in areas like design where physical activity is preceded by cognitive and reflective activities. For example, if we start with a simple picture of a man making an object, then it would be reasonable to argue that in as much as he has a definite cognitive anticipation of the probable object (that is, he is not simply experimenting by trial and error with the latencies of his tools and raw materials) then he is acting analogously to a designer as well as being a maker. His cognitive anticipation of the object is part of the field of tools and raw materials that constitute his 'instrumental set'. Design, as we know it, can be seen as the socially differentiated transformation of the reflexive cognition of the maker in terms of the latent possibilities of his tools, materials and object types. Its object is not the building, but at one remove, sets of instructions for building. The activity called architectural research can be derived by an exactly similar transformation, namely a socially differentiated transformation of the reflexiveness of the activity of design upon itself, that is, its object is design, and its product takes the form of rules or rule-like systems for design which stand in the same relation to design as design does to building. As in other sciences, it finds the best way of doing this is by addressing most of its strategies to the 'real' world, and if we are not careful, this, coupled to the fact that the activity is necessarily multi-disciplinary, tends to conceal the 'deep structure' of the activity. This is why we suggested earlier that we should call the research activity meta-design. At least this might begin to emancipate us from the silly (but pervasive) idea that the outcome of research is 'knowledge', to be contrasted with the absence of such 'knowledge' in design.

We can perhaps clarify the characteristics of design as a cognitive activity by going back to the very simplified situation we have just referred to, to see if we can discover what there is in the maker/designer's field, and go from there to see how it differs today. Here we owe some debt to Lévi-Strauss's discussion of 'bricolage' as an analogy to myth-making.[15]

We can imagine a man and an object he will create as though separated by a space which is filled, on the one hand, with tools and raw materials which we can call his 'instrumental set', (or perhaps technological means) and on the other, a productive sequence or process by which an object may be realised. If time is excluded from the space, we can conceive of the 'instrumental set' as though laid out on a table, and constituting a field of latencies and pre-constraints.[16] If time is in the space, then the instrumental set is, as it were, arranged in a procedure or process.

The total field thus exhibits two types of complexity, and we may allow that the maker is capable of reflexively making both types of complexity (the latencies of the instrumental set, and the distribution in process-time) the objects of his attention.

Two basic strategies appear to be open to him. He can either distribute the latencies of the instrumental set in process-time according to some definite cognitive anticipation of the object he is creating, that is, pursue a definite design or plan, which may be based on an analogy or on pure imagination, as it may be conceived in terms of the familiar products of the instrumental set. Or he can, as it were, interrogate his instrumental set by an understanding of its latencies in relation to general object types. In both strategies, an understanding of the latencies of instrumental sets and a knowledge of solution types is of fundamental importance. In other words, the maker's capability in pre-structuring the problem is the very basis of his skill, even if he wishes to proceed heuristically by interrogating his instrumental set and exploring unknown possibilities by a dialectic between his understanding of the latencies and limitations of the instrument set and his knowledge of solution types. On this basis we would argue that design is *essentially* a matter of pre-structuring problems either by a knowledge of solution types or by a knowledge of the latencies of the instrumental set in relation to solution types, and that this is why the process of design is resistant to the inductive-empiricist rationality so common in the field. A complete account of the designer's operations during design would still not tell us where the solution came from.

But there is an escape clause. As with science, it is not a matter of *whether* the problem is pre-structured but *how* it is pre-structured, and whether the designer is prepared to make this pre-structuring the object of his critical attention. From here we would go on to suggest that the polarisation we have assumed between rational and intuitive design should be reformulated as a polarity between reflexive design (that is, design which criticises its understanding of the latencies of instrumental sets and solution types) and non-reflexive design (that is, design which is simply oriented towards a problem and which therefore operates within the known constraints and limits of instrumental sets and solution types). To equate rationality with a certain type of systematic procedure appears therefore, quite simply, as a mistake.

Part 2

It is obvious that today the designer operates in a field which is considerably more complicated than the one we have described, based on a man making an object. The notion of pre-structuring is *necessary* to any conceptualisation of design, but not *sufficient* in itself. We have to look at the complications and how they have evolved, in order to complete our conceptualisation of the designer's field and his operations in it.

The most obvious difference is that design is not simply the reflexive/cognitive aspect of making an object, but a separate, socially differentiated activity with its own internal dynamic and its own end product, namely sets of rules for making artefacts. It is also a highly specialised activity, carried out by a clearly defined social group. There is therefore no direct link between interrogating the instrumental set and the result as it is likely to be experienced by those who use it. We thus require a great deal of information about the latter in order to interrogate the instrumental set.

We can explore the consequences of this development by trying to imagine what life was like when we had designers, but not user-requirement studies. How did we live without them? The answer seems quite simple. Notions about the user were built into the instrumental set and the solution types. The instrumental set was comparatively unsophisticated and had in any case been developed mutatively over a long period. It was already an expression of the basic physiological requirements of users in terms of available technology, and probably a reasonable approximation of their psychological and other expectations. The solution types had been similarly evolved, and contained already the notions of use and activities within the building. We could say that contained in the instrumental set and the solution types was an implicit, historically evolved code, which linked the means to the ends. It would be difficult to decipher and reconstruct, but we can see that it was there, and, in principle, how it got there.

Since those days we have seen developments like the proliferation of building types, and the proliferation of instrumental sets (technological means) and a formal organisation of the process which results in most activity being of a one-off kind with the simple effect that the users' needs in terms of activities, physiological requirements and cultural expectations are no longer contained, as it were, in the instrumental sets and solution types. A much freer, more indeterminate situation appears to exist. This deficiency is made up in terms of information which is expressed in terms of the users rather than in terms of buildings, and the designer operates a kind of *informal code* for linking one to the other. Part of the outcome of research in the past has been a piecemeal and atomistic, partial replacement of the codes, by formal rules which, when implemented, often have the unfortunate effect of dictating the whole design (the 2 per cent daylight factor is a classic example). The designer's task becomes something like the utilisation of these codes in order to link the information he gathers about the project to his interrogation of the increasingly prolific instrumental sets, or his manipulation of solution types. He has to deal similarly with the proliferation of information extraneous to the particular problem relating to standards, constraints, quantification rules and so on. In this situation, it is perhaps no wonder that the designer (unless his ambitions are, frankly, artistic) welcomes the prospect of a logic whereby solutions can be synthesised out of information. It offers him the prospect of eventual escape from the contradiction of actually working by the interrogation of instrumental sets or the adaptation of solution types, as he always did, but being expected to utilise a procedure of optimising information which bears little relation to building, except where piecemeal, atomistic rules have been developed. Perhaps we should add one more point to this analysis: that the informal codes the designer must use to link information to build outcomes are also instances of problem pre-structuring.

If this is a reasonable characterisation of the principal elements in the designer's field, then at least we are some way to understanding why designers do not produce better buildings out of the information research provides, and why, with expanding technological means and user requirements, the theoretical open-endedness of architectural problems leads to so little fundamental variety in the solutions proposed. With a proliferation of poorly understood instrumental sets, increasingly masked by unrelated information, we would expect that a retreat to the most basic form of pre-structuring – the adaptation of previous solutions – would become the only viable way through the morass. Far from helping the designer escape from his preconception, the effect of proliferating technology and information is to force the designer into a greater dependence on them. Innovation becomes more rather than less difficult, but the diffusion of uncritical innovation would become more rapid. A situation develops in which a few experiment and others adapt solution types, without understanding or evaluating the rationale of the original experiment. The net result is unstructured innovation, with slow and piecemeal feedback, giving the impression of arbitrary shifts in fashion. This seems a not unreasonable account of the situation we have, and would explain why even well-disseminated and well-presented information – such as widely exists in the UK from the Building Research Station and government departments – either does not lead to an improvement in the product or does so only in a haphazard way.

We would also suggest that this leads to a situation in which students are learning two different and largely unrelated strategies: methods of analysing a problem into its elements; and a knowledge of informal codes and solution typologies, which they pick up almost as by-products of architectural education, and which act as the pre-structuring that enables them actually to design buildings.

Part 3

We have argued that the chief elements present in the designer's field are *knowledge of instrumental sets*, *knowledge of solution types*, *informal codes* and *information*. These cannot usefully be reduced to homogenised 'information', although it is possible at a theoretical and formalised level. Now we would like to use these ideas to try to construct a lifelike conceptualisation of design as an activity.

These elements constitute the designer's field, his set of latencies and pre-constraints. Somehow these are to be distributed in a process unfolding over time. We will need to introduce one or two further basic ideas as we proceed, but we hope that these will either be from those we have already discussed, or simple logical statements of an unproblematical kind.

For example, it seems unproblematic to say that when a design problem is stated there are, theoretically at least, a number of solutions open, probably a very large number. Yet only one of these possible solutions will be the final one that is built. We may reasonably say that some process of *variety reduction* has taken place. The variety of possible solutions has been reduced to one unique solution by some means. The succession of documents produced during design reflects this progressive reduction of variety. More and more specific drawings, for example, exclude more and more detailed design possibilities. We would like to introduce this as a basic idea in our conceptualisation of design.

A second idea we would like to introduce is that of *conjecture*. Here we would like to go back to science. It was once thought that conjecture would have no place in a rigorous scientific method. It was thought to be akin to speculation, and science sought to define itself in contradistinction to such notions. Since Popper, we know that science cannot progress without conjecture; in fact, together with rigorous means of testing, conjectures constitute the lifeblood of science. Conjectures come from anywhere, and because they are not derived from the data by induction, it does not mean that the process of thought of which they form part is any the less rational or rigorous. What is irrational is to exclude conjecture. So, we will include it in design.

How does the reduction of variety from many possible to one actual solution take place? Obviously, anything we can say here will only be an approximation of any particular case. But our aim is to try to understand the process of design as it exists in the real world, in order to try to define the contribution of meta-design. What we are aiming at is some more or less true to life approximation of the psychology of design, bearing in mind that design is a practical as well as a cognitive activity, and that design problems do not happen in a social vacuum, but are socially constructed.

Beginning with a theoretically open problem, with an unlimited number of solutions, it should be clear that the variety of possible solutions is already reduced before any conscious act of designing begins by two sets of limiting factors, one set external to the designer, the other internal. The first set we can call 'external, variety-reducing constraints' and these can often be quite powerful, or even totally deterministic of the design. For example, a client who says categorically, 'I want one like that' has already reduced the number of possible solutions to one. More often the external constraints will be of a less overt, but still powerful, kind, such as norms of appearance, availability of technological means, costs, standards and so on. Some of these will not be fully understood by the designer at the outset, but as he specifies them, their role as variety reducers will become clearer.

The second set we can call the 'internal variety reducers' and these are an expression of the designer's cognitive map, in particular his understanding of instrumental sets and solution types. This notion of the pre-existing cognitive map is very important indeed, because it is largely through the existence of such maps that any cognitive problem-solving activity can take place. They are, and must be, used by the problem-solver in order to structure the problem in terms in which he can solve it. It acts as a kind of plan for finding a route through problem material that would otherwise appear undifferentiated and amorphous. Its role is equivalent to the role of theory and theoretical frameworks in science. Data are not collected at random. What is to be called data is already determined by some prior theoretical or quasi-theoretical exercise, implicit or explicit.

We have to recognise, therefore, that before the problem is further specified by the gathering of data about the problem, it is already powerfully constructed by two sets of limiting factors: the external constraints (although some of these may still be poorly understood) and the designer's cognitive capability in relation to that type of problem. It is quite likely that these latent limitations are already being explored right from the beginning, if the designer is conjecturing possible solutions, or at least approximations of solutions, in order to structure his understanding of the problem, and to test out its resistances. There is also a very practical reason why conjectures of approximate solutions should come early on. This is that a vast variety of design decisions cannot be taken – particularly those which involve other contributors – before the solution, in principle, is known.

As the designer collects and organises the problem data, and data about constraints, his conjectures acquire sharper definition. Previously he was not able to test them out in a very specific way. Now he has an increasing fund of information against which to test them. He will also be using this information heuristically by using it in relation to his informal codes (see above) by which abstract requirements are linked to built outcomes, and conjecturing further specifications within his roughly conjectured solutions. Information which has been used heuristically can also be used to test the new conjectures. Conjecture and problem specification thus proceed side by side rather than in sequence. Moreover, conjectures do not, on the whole, arise out of the information although it may contribute heuristically. By and large, they come from the pre-existing cognitive capability – knowledge of the instrumental sets, solution types and informal codes, and occasionally from right outside – an analogy perhaps, or a metaphor, or simply what is called inspiration. At least, within this conceptualisation of design, we do not have to say that designers who use these last three types of source for conjecture are acting in a way that is markedly different from the architect with more modest ambitions. He has simply widened the scope of his conjectural field, sometimes moving right beyond the limits of the instrumental sets that are available.

When a conjectural approximation of a solution stands up to the test of the increasingly specific problem data (bearing in mind that it is always possible to collect more data and to produce more conjectures) a halt is called to both conjecturing and data gathering, and a solution, in principle, is agreed to exist. Further specification then takes place (that is, further variety reduction) by completing a full design, and this is followed by a further refinement when the final production drawings are made. Unless the designer has great foresight, it is likely that further refinements will be made at the building stage.

We believe that this is, more or less, how design happens in most situations, and we believe, moreover, that it is as rational a process as is possible in the complex circumstances, not sub-rational because it is not 'systematic' and because so much depends on how the designer pre-structures the problem. This outline model differs from the analysis-synthesis model (which we take to be the dominant notion in design method studies, hitherto) in several important ways. First, its core stratagem is conjecture-analysis rather than analysis-synthesis. Second, the purpose of analysis is primarily to test conjectures rather than to optimise by logical or magical procedures. The notion of optimising, which architects believe they carry out, can be easily contained within a conjecture-test psychology of design. Third, the solution, in principle, is allowed to exist at a much earlier stage than in the analysis-synthesis model. Fourth, the model shows the path of convergence on a unique solution without introducing notions like the optimisation of information which, while attractive theoretically, are largely unlifelike and unworkable. Fifth, the model suggests, *within its basic concepts*, the possible origins of solutions in principle, a matter on which the design methodologists are notoriously silent or mysterious. Sixth, the model corresponds to the observed sequences of products of design, namely a set of descriptive documents of increasing refinement and specificity. Seventh, it recognises implicitly that both information and conjectured solutions are inherently incomplete, but a stop has to be called somewhere. This is precisely equivalent to the situation in science. Eighth, and perhaps most important, the model emphasises the importance of the designer's pre-structuring of the problem, rather than denigrating it. It recognises that architects approach – and should approach – design holistically and not piecemeal.

Part 4

What does this have to say about research? We have already argued that presenting the 'results' of research in the form of packaged information or quantification tools does not seem to lead easily to better solutions. Perhaps the model will help to explain why. It is largely because, unless research can influence designers at the stage of pre-structuring the problem in order to understand it, then its influence on design will remain limited.

To explore this further, we might usefully examine the outcomes of research in terms of the four main types of elements which characterise the designer's field, namely instrumental sets, solution types, codes and information. It can be seen that much research of a purely technological kind (still by far the largest investment in building research) has its outcomes in terms of instrumental sets. Development work extends this into solution types by proposing exemplars. Research which aims to provide a method of checking design proposals against abstract requirements can be seen as a partial formalisation of codes (partial because it is concerned with testing rather than generation and it is piecemeal). And research which has its outcome in the form of 'results', rather than a tool, falls into the field of information.

It can easily be seen that the first and last of these do not really help the designer to design. They normally increase the complication of the field and obscure its structure. Certainly, they do not help the designer much at the stage of pre-structuring the problem, and if they do so, it can only be in a haphazard way. The exemplars and prototypes that are the outcomes of development work certainly help the designer to pre-structure his problem, but only if he proceeds in a largely imitative way. If the development is inadequate in any respect, it leads to a proliferation of these inadequacies.

Over and above this, the prototype may be poorly understood, or badly adapted. Research in the third category is similarly unhelpful at the crucial stage of pre-structuring. It may provide a means of eliminating errors at the design testing stage, given that the designer is able to use them properly, but we can hardly conceive of the designer being able to effectively utilise the full panoply of such techniques that would be required to cover all aspects of the design.

Of the four, only the development model can demonstrate, to the designer, new ways of pre-structuring his problem. In spite of its disadvantages, its potential usefulness should not be underestimated. We could say that it suggests an organisational solution to the problem of linking research effectively with design. If research workers work with designers in producing experimental prototype solutions, which are intensively monitored and improved, then explained and publicised, then research itself benefits by becoming part of a dynamic process from which it can continuously learn and develop its concepts. In the past, development work in building has tended to lack both the deep involvement of research workers and a properly developed monitoring function linked to a building programme. If both of these are provided for, there is at least an opportunity for sustained development over a period. By the quality and conviction of its exemplars, it can lead quite rapidly to a diffusion of real improvements in solutions.

On the other hand, the disadvantages of relying wholly on this fail-safe means of linking research with design are strong. The individual designer becomes severely constricted, problems of poor interpretation and debasement are likely to arise, creative innovation may be cut off or inhibited. Is there not some way in which research may help the designer to pre-structure his problems more effectively without predetermining the solutions?

We believe there is, and that it lies in the notion of codes, the third element in the designer's field. Informal or implicit codes, we suggested, were used by the designer to link abstract functional requirements with instrumental sets, which no longer contained such codes. Taken together as a system, they constitute a kind of quasi-theory by which the designer structures his problem and finds a route through it – or through as much as is left of the problem after other external and internal constraints, including solution types, have had their say. Sometimes these codes are formalised and externalised in a rather pragmatic and programmatic way as 'architectural theory'. The influence and rate of diffusion of such externalisations is often very considerable.[17] On occasion their impact is such as to have a marked effect on the development of instrumental sets.

The idea we are working towards, stated simply, is that research should aim (and is already beginning to aim) at the progressive reconstitution of the codes on a conceptual base by studies of people and their built environment which are oriented towards theory rather than 'results'. This is a complex and long-term aim, but it is entirely consistent with the normal impact of scientific work on human activities. The difference between a craft and a technology is not research results, but theory which brings structure and classification into phenomena, and allows the possible to emerge from an understanding of the actual. In any problem-solving activity, theory is the essential link between science and action. Without theory and its classificatory and route-finding possibilities, design is likely to remain, even in a field of endlessly proliferating scientific 'information', a kind of craft without continuity.

Here we come back to the reasons for optimism about architectural research. It seems to us that we are seeing the development of strong research programmes, which are architectural in that they deal with broad bands of connected factors in design, and fundamental, in that they are concerned with theories which actually relate to these levels of integration, rather than theories about isolated factors in environment. We would therefore like to try to explain what we see as the emerging structure of architectural research, why it is theoretical in a design as well as a scientific sense, and why it appears capable in the long run of affecting the ways in which problems are pre-structured by designers.

We can best explain this by asking a question. What, in theoretical terms, is a building? On the grounds that buildings are not gratuitous but entirely purposeful objects, we would define a building as a realisation of a number of social functions with an effect of ecological displacement. By specifying these functions and displacement effects in sufficiently abstract terms, we can formulate an adequate theoretical description of what a building is (such that anything which lacks one of them is not a building, and that if an object is a building, it will fulfil all these functions whether by intention or as a by-product) and what its displacement effect is in terms of a four-function model. These are not true for all time, but are a historically accumulative set which define more or less what a building is at this point in time.

First, a building is a climate modifier, and within this broad concept it acts as a complex environmental filter between inside and outside, it has a displacement effect on external climate and ecology and it modifies, by increasing, decreasing, and specifying, the sensory inputs into the human organism.

Second, a building is a container of activities, and within this it both inhibits and facilitates activities, perhaps occasionally prompting them or determining them. It also locates behaviour, and in this sense can be seen as a modification of the total behaviour of society.

Third, a building is a symbolic and cultural object, not simply in terms of the intentions of the designer, but also in terms of the cognitive sets of those who encounter it. It has a similar displacement effect on the culture of society. We should note that a negatively cultural building is just as powerful a symbolic object as a positively (that is, intentionally) cultural one. Fourth, a building is an addition of value to raw materials (like all productive processes), and within this it is a capital investment, a maximisation of scarce resources of material and manpower, and a use of resources over time. In the broader context of society, it can be seen as a resource modifier.

In brief, a building is a *climate modifier*, *a behaviour modifier*, *a cultural modifier* and a *resource modifier*, the notion of 'modification' containing both the functional and displacement aspects.

Each of these functions can be conceived of separately as a people–thing relationship and each, in contrast to research oriented towards the 'atoms of environment', deals with a holistic set which constitutes *one way* of looking at a design problem. Each is capable of developing theory about people and their built environment. We would argue that research is gradually organising itself within these foci as a set of interdependent, theory-oriented and largely structural studies, and that these are emerging as the fundamental disciplines of architectural research, and providing the base within which various disciplines become integrated and lose their identity.

Part 5

It is notable, by the way, that the emphases implicit in this model shift architectural research right away from the study of procedures of design and into the study of buildings and their occupants, as well as away from 'results' and towards theory. We are beginning to look again at ends rather than means.

How will such research contribute to design? We have argued at the general level that it will progressively enable us to reconstitute codes from a theoretical base concerned with the relations between physical environments and those who experience them. We may add, first, that we conceive of this happening not in a positivistic and piecemeal way, but, because of the theoretical base, in a more holistic, non-deterministic and heuristic way. But this is too general a statement to be useful. We must specify further what we mean, and show why we can use this idea to escape from the idea of once-and-for-all 'knowledge' and allow for fundamental shifts in the theoretical bases by which we define 'knowledge' which will undoubtedly occur. We appeal again to the lessons of science.

In spite of periodic epistemological crises, paradigm switches and the progression and degeneration of research programmes, science continues to build its usefulness (as it has always done) on the strength of precise descriptions of the world. The theories on which these precise descriptions are based may be incomplete and even wrong, but they enable us to organise more and more of the world into useful cognitive schema which, among other outcomes, enable us to conceive the possible out of a study of the actual. It might not be going too far to characterise the history of science as a series of immensely fertile delusions.

We do not therefore need to invoke the idea of 'knowledge' in order to propose that out of the notion of a building as a multifunctional object, and design as a multi-theoretical activity, we can begin to build up theory-based descriptions of the basic elements in design. These basic elements include ranges of activities, movements, perception-motivated actions, social intercourse patterns, spaces and the environmental criteria that will satisfy a classified range of possible uses, coded and described in terms of the technologies which make them possible. Such a breakdown we might call a *base component classification* for environmental action, which would shift both in response to theoretical changes and also in response to changes in the environmental objectives of society. From the point of view of the designer, such classifications and code formalisations would not be deterministic, or constitute a set to be specified in relation to problem information, but would constitute an extension of the designer's basic cognitive capability, and provide him with – and this is really the point about science – a *position of strength from which to make his conjectures*. In other words, he would be using theories operationalised and specified, as far as possible, in terms of externalised codes, linking instrumental sets to human usage, as a basis for proposing his own further modifications to the environmental field.

The implications of this for the current formal structure of design activity – particularly those concerned with briefing, one-off user studies and the designer's ability to reinterpret the 'client's requirements' – are enormous, and to examine them in detail would require another paper. To give one example, in the area of activity-space relations, we can foresee the possibility of moving from the 'activity-space fit' notion, which is implicit in current practice, towards much more fundamental theories about the capability of certain types and configurations of space to contain an unpredictable variety of activities, perhaps with consequences for the idea of building types, and even for the size of cities. Such theories are not pseudo-deterministic ways of telling the designer what the outcome of his design will be, but strong and cumulatively developing bases for conjecturing possible futures.

If we are right in thinking that this is the underlying direction of the new lines in environmental research, then the notion of research simply as a service to design and the by-product of an eclectic variety of disciplines has to go by the board. Research is, of course, necessarily multi-disciplinary. In fact, in the environmental field there appear to be no limits to the disciplines that could contribute to the advancement of the subject. But the contributions of the wider areas of science will only become effective through the integrative theories which will increasingly form the fundamental disciplines of environmental action itself, and these disciplines are not separate from design but extensions of it in that their subject matter is design just as the subject matter of design is sets of instructions for building.

This is not a strange or unique arrangement. In fact, it is very similar to science itself, seen in its broadest terms as one of the activities of society. Through science we continuously modify the world we live in and our understanding of it – that world and that understanding that it is the aim of science to study.

Notes

1. Editors' note: this article was first drafted by Bill Hillier working with Adrian Leaman – indeed, its contents resonate strongly with other papers by Hillier and Leaman, including those in this volume. The article was presented as a paper at the 3rd EDRA conference held in Los Angeles in 1982, which was attended by Bill Hillier, John Musgrove and Pat O'Sullivan, but not by Adrian Leaman. The paper was published in the EDRA proceedings with Bill Hillier, John Musgrove and Pat O'Sullivan listed as co-authors, and has been republished with the same co-authorship attribution (for example: B. Hillier, J. Musgrove and P. O'Sullivan, 'Knowledge and design', in *Developments in design methodology*, edited by N. Cross (New York: John Wiley & Sons, 1984), pp. 245–264. We note that in a later paper (J. Musgrove, 'Educating environmentalists', *Transactions of the Bartlett Society* 9 (1973): 20–35), John Musgrove endorses the 'conjectures and analysis' model of design advanced in 'Knowledge and design' and refers to it in two different ways: as Hillier, 1972; and also as Hillier, Musgrove and O'Sullivan, 1972. We are not aware of any paper by Pat O'Sullivan addressing issues related to this paper. Here we add Adrian Leaman as a co-author, thus correcting a historical accident (eds.).
2. Editors' note: the 1984 republication of this paper added a hyphen to this neologism. We have followed this revision to assist in comprehension.
3. In this paper the words 'architecture' and 'architectural' are used as shorthand for the built environment and its action systems, as a sub-system of environmental action and modification as a whole. It is not intended to refer to the activities or ideology of a particular professional group. We apologise for any confusion this may cause.
4. The analysis in the early part of the paper refers principally to the UK situation as it has developed in the last 15 years or so, but we hope that the arguments will retain most of their validity when applied elsewhere.
5. Royal Institute of British Architects Research Committee, 'Strategies for architectural research: Architectural research and teaching', (1970): 3–5.
6. For example, A. Colquhoun, 'Typology and design method', *Arena: Journal of the Architectural Association* 83 (1967): 11–14.
7. S. C. Keene, *Introduction to meta-mathematics* (Amsterdam: North-Holland Publishing Co., 1959); E. Nagel and J. R. Newman, *Gödel's proof* (London: Routledge and Kegan Paul, 1959).
8. J. Piaget, *Structuralism*, translated [from the French] and edited by C. Maschler (London: Routledge and Kegan Paul, 1971).
9. Piaget, *Structuralism*.
10. K. Popper, *The logic of scientific discovery* (London: Hutchinson, 1959); K. Popper, *Conjectures and refutations: The growth of scientific knowledge* (London: Routledge and Kegan Paul, 1963).
11. T. S. Kuhn, *The structure of scientific revolutions* (Chicago: University of Chicago Press, 1962).
12. I. Lakatos and A. Musgrave, eds, *Criticism and the growth of knowledge* (Cambridge: Cambridge University Press, 1970); also, I. Lakatos, 'History of science and its rational reconstructions', in *PSA 1970: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, edited by R. C. Buck and R. S. Cohen (Dordrecht: Reidel, 1970), pp. 91–136. Eds note: this was cited as 'forthcoming, in mimeo' when 'Knowledge and design' was published in 1972.
13. G. Kelly, *The psychology of personal constructs* (New York: Wiley, 1964).
14. Editors' note: this is now known as the BRE (Building Research Establishment) (eds).
15. C. Lévi-Strauss, *The savage mind* (London: Weidenfeld and Nicholson, 1962) (English edition, 1966).
16. Hyphen added for clarity. The editors point out that 'latency' at the time of writing referred to a potentiality, rather than the more common usage nowadays to refer to lag, or delay.
17. Le Corbusier, *Towards a new architecture* (London: Architectural Press, 1946) (First English edition, 1927).

2 The man–environment paradigm and its paradoxes (1973)

Introduction to 'The man–environment paradigm and its paradoxes'

Alan Penn

Re-reading 'The man–environment paradigm and its paradoxes', I am struck that it is a paper both of and ahead of its time. Since its publication, much of sociology has passed through a 'spatial turn' and the argument against naive determinism seems outdated. In biology, just 20 years after DNA's structure was determined, there was optimism that all of life's secrets would be revealed. Today there is a creeping realisation that all is not so easy, while the contemporary critique of neo-Darwinism through a 'niche construction' lens[1] is creating a more nuanced understanding of organism–environment relations. In philosophy, the first English translation in 1962 of Martin Heidegger's 1927 *Being and time*[2] was greeted with derision by the anglophone philosophical establishment. I have little doubt that Bill Hillier was sceptical of its substance, but its key contribution – that ontologically cognitive organisms are born into and inseparable from their environment – directly aligns with Hillier and Leaman's thinking.

At the time of writing, both authors were working at the RIBA Intelligence Unit developing the architectural profession's response to the 1958 Oxford Conference on Architectural Education at which Richard Llewelyn-Davies had outlined a scientific approach to the education of an architect[3] that triggered a student revolt at UCL's century-old 'Beaux Arts' School of Architecture. By 1970, as Head of the Bartlett, he had brought together the schools of architecture and planning in a renamed 'School of Environmental Studies', and radically remodelled the curriculum by appointing professors and establishing research programmes in a range of disciplines from building physics and psychology to construction and ergonomics, which he held formed the 'meeting grounds' of knowledge essential to architectural practice.

In 1972 Hillier and Leaman were invited to establish a new MSc degree in 'environmental theory' at UCL. The title and focus of this paper are therefore no coincidence, forming both a critique of Llewelyn-Davies' meeting grounds philosophy (in which architecture itself was notable by its absence) and a manifesto for both the master's degree and for the programme of research that Hillier and

Leaman initiated and Hillier pursued in subsequent years when he became Director of the University College London Unit for Architectural Studies.

One of Hillier's research maxims was that a paradox indicates an underlying problem with the paradigm assumptions in a field. Michel Foucault's *Order of things*,[4] published in English in 1970, framed the concept of the 'episteme' charting paradigm shifts from classical times to modernity, while Thomas Kuhn had showed scientific progress rests on shifting paradigms.[5] Sociology at the time was insistent that the environment cannot have a determining effect on human action if free will is to be maintained, and yet architects were certain that their work mattered while residents of failing public housing projects were equally certain that their environment was at fault. However, many failing estates were a product of Modernist functional design. For example, Alison Ravetz's study of the notorious Quarry Hill estate in Leeds[6] charted its origins in European Modernism through to its demolition in 1974. To the sceptical sociologist, functionalism's failure served merely to support their thesis that the environment could play no active role.

Hillier and Leaman's resolution came in relational and logical parts. Jean Piaget's early studies of child development[7] showed that knowledge is constructed through interactions between individuals and their environment, while his later structuralist work traces the origin of knowledge in the empirical world, theorising that resulting logical structures form a basis for rationality. From Ferdinand de Saussure[8] they took the notion that in systems of cultural transmission, the physical utterances of speech (*parole*) are drawn from a logical field of possibility in language (*langue*). It is the field of logical possibility that allows for language to be both intelligible and to be used creatively to say things never before said.[9] The insight here, that spatial relations are fundamentally logical, led to Hillier and Leaman's development of the idea of 'morphic language' in their 1976 paper 'Space syntax', and ultimately to Hillier and Hanson's *The social logic of space*.

The latter marked not only the end of Hillier and Leaman's creative partnership, but also a shift into analytic research. If spatial relations in our constructed environment are produced through social processes, then by analysing the spatial object it should be possible to reveal something about those processes. The development of spatial analysis marked the empirical breakthrough in the field, including the successful identification of the consequences of design on human movement.[10] This put flesh on the bones of the emerging social theory by providing a mechanism through which architecture could feed back into society. It also made possible a means to intervene in practice by analysing design options. However, it must be remembered that in 1973 these empirical steps had yet to be taken. The authors were at a hypothesis-generating stage in their quest to understand architecture.

The subsequent development of space syntax research and practice has been criticised for being overly mechanistic and reductive. This is unsurprising – if one identifies mechanisms through which society forms and transforms itself over time, surely that is mechanistic? The focus on space to the exclusion of institutional and cultural strata in society is no doubt reductive. However, mechanisms do not imply deterministic mechanical processes – and the processes at play here must largely be probabilistic. Similarly, the notions that were to become central to the emerging theory such as description retrieval, the virtual community and the movement economy imply no single direction to causality, but that feed-forward and feedback are equally involved. The processes are probabilistic, emergent, dynamic, and while in an essential sense bi-directional, they are not conflationary.

I am with Richard Wilmott in his argument[11] that for the purposes of analysis, it is perfectly reasonable to separate the social world into different strata. The organisms we see in the world *are* distinct from each other and from their environment. As cognitive entities they each seek to make sense of their surroundings and equally are constrained by these. The child develops by learning in their own body and in the world. But for learning to be possible the world must be 'learnable' and display enough regularity of association between its different strata to be intelligible. In scientific analysis, however, we gain much by looking at the different strata: space, society and so on separately before reintegrating them. Central to reintegration lies a quest to identify the underlying logical structures of Popper's 3rd world and Piaget's structuralism that make intelligibility possible. I suspect that this was central to Hillier's programme of research in subsequent years, for which 'The man–environment paradigm and its paradoxes' provided a manifesto.

The man–environment paradigm and its paradoxes

Bill Hillier and Adrian Leaman

Originally published: Hillier, B. and Leaman, A. 1973. 'The man–environment paradigm and its paradoxes'. *Architectural Design*, 43, 507–511.

A paradigm is any structure of ideas, scientific and philosophical, that we take for granted in order to do research. Because we must take some such structure for granted, paradigmatic ideas tend to become invisible. We forget that they are there, and regard them as natural. Only occasionally are they brought into question and perhaps replaced. The related notions of 'organism' and 'environment' constitute such a paradigm base. Translated into the sciences of man they become the 'man–environment paradigm', which is much more problematic than it appears at first sight. In fact, it contains anomalies, which appear in research as unresolvable paradoxes. In a profound sense, the concept of environment itself obstructs the development of a relevant theory for environmental analysis.

The concept of environment

In science, in theories of knowledge and in everyday language, the notion of 'environment' is pervasive. It appears both obvious and necessary, so much so that it is never directly questioned. It appears to be a natural dimension of our way of thought.

But like many natural seeming ideas, closer inspection reveals the artifice and the cracks. In the first place, the concept turns out to be an unexpectedly recent one, as far as science is concerned, dating only from the later part of the eighteenth century. Secondly, the precise meaning of the term is very difficult to pin down in any scientific sense. Like other powerful concepts, it is prolific with meaning and suggestion. The term has acted as a 'basin of attractors', drawing other meanings and concepts into itself, as a river draws tributaries as it defines, and is defined by, the landscape through which it passes.

These considerations suggest that we should at least look at the possibility that the idea of 'environment' is an element in some epistemological formation, or paradigm, rather than an inevitable aspect of consciousness. The history of the concept suggests not only that this is so, but that its paradigmatic influence is even more far reaching than appears at first sight, particularly in the 'social and human sciences'.

Like most scientific terms, the concept of environment does not exist in isolation. It is part of a complex of ideas. Its most immediate neighbour is the concept of the 'organism'. Together they constitute the paradigmatic notion that realities involving living structures could best be looked at in terms of the relation between 'organism and environment'. This set of ideas, which appear originally in biology (as we would now call it) in the latter part of the eighteenth century, has a clear development prior to its emergence in recognisable form.

According to Canguilhem, the concept 'environment' ('milieu' in the French original) came into use as a result of the combination of two earlier concepts, one from mechanics, the other from 'anthropo-geography'.[12] The idea from mechanics was Newton's solution to the physical problem of action at a distance, within the assumption of a mechanistic 'cause and effect' universe. 'To suppose that one body may act on another at a distance, through a vacuum, without the mediation of anything else, is to me so great an absurdity that I believe no man, who has in philosophical matters a complete faculty for thinking, can ever fall into.'[13] How could action at a distance occur if space was empty? Newton's solution to the problem posed by his own physical theory (but not by previous ones) was the concept of the 'ether'. This was a pervasive medium present throughout the universe in which forces could be transmitted. Newton himself applied this concept to biology in order to link the mechanical propagation of light outside the organism with the reactive physiological events that took place within the organism. The 'ether' pervaded the earth, the atmosphere, the eye, the nerves and the muscles, enabling continuous chains of action to take place on the basis of physical causality.[14]

The idea from 'anthropo-geography' was that the variability of the human species could be accounted for in terms of regional variations, in particular climatic and other geographic factors. This idea began to acquire scientific status towards the end of the eighteenth century and remained in a variety of forms a major epistemological concept in geography until the post-war period.

Both ideas are even older in more intuitive versions. The first echoes the ancient notion of each species in its 'proper element'. The second recalls older notions of climatic determinism, as found for example in the work of lbn Khaldun, the fourteenth-century Arab historian. But we are here interested in the admission of the concept 'environment' to the field of legitimate science and this appears to have initially taken place in the work of Buffon.[15] Buffon attempted to show how the variations of animal species could be explained by environmental influence on organic form and behaviour. This scientific formulation was a major event in the history of biology and, conceptually at least, was a major step away from fixed species taxonomy towards evolutionism. The concept was developed into the first primitive version of an evolutionary theory of the animal world by Buffon's pupil, Lamarck. Lamarck thought that the environment could account for the species by looking at the functions of the organism in relation to its environment. This implied also that what was important to our understanding of the organism was not the classification of visible differences but the organic structure of the whole functioning organism. Thus, the concepts 'environment' and 'organism' took their place in natural science side by side, with the notion of function playing a catalytic role. This is an early scientific version of the 'man–environment paradigm'.

But such a powerful and pervasive idea as 'man–environment' does not spring from a single source or follow a single line of development. Scarcely less important is the philosophical history; and even more important, the influence of scientific ideas on philosophy. This crystallises in the effort of Kant to find a new path through the old paradoxes of rationalism (inborn ideas) and empiricism (acquired ideas – the child's mind is a blank sheet to be written on). To understand this, we have looked back to Descartes and his partition of the universe into thought and matter, and his consequent mechanistic treatment of physical realities, organic and inorganic. English empiricism was derived from this by Locke through a simple extension: man as a thinking being was the sum of inputs into the organism, structured by the 'laws of association', to form the patterns we call 'thought'. This account of man was on the whole accepted in the eighteenth century as part of the package deal scientific-paradigm-cum-philosophy created by Cartesianism and refined by Lockian empiricism and Newtonian science.

But the very success of Newtonian science created a paradox in the form of a contradiction between scientific and psychological empiricism. If Newtonian science was the model of 'scientific certainty', as was widely supposed, how could this be accounted for through the empiricist model of man which appeared to imply a large measure of contingency in the concept of 'knowledge'? How, in short, could the 'association' of inputs into an organism, which as a result of these inputs could think, account for certainty such as Newton appeared to have achieved?

Kant's solution was that knowledge is possible because certain empirical concepts like space, time and causality are built into the structures of our perception and understanding in such a way that we cannot escape from them. Although this basic Kantian concept was quickly shown to be mistaken (by, for example, the discovery of non-Euclidean geometries), the Kantian synthesis is of the highest importance in the history of philosophy and meta-scientific thought. It constructed again an interdependence between man and his environment, replacing the 'preestablished harmony' implicit in the rationalist programme of Descartes and Leibniz as the basic notion for our scientific understanding of the world, and constructed it for purely logical reasons. Yet the outcome is congruent with the scientific concept of organism–environment as it evolved in eighteenth-century biology. This is evidence of a paradigm in formation.

A further influence is also important: the effect of the scientific concept of organism–environment on the concepts of man and society. The organism–environment concept was the means by which time, change and history were introduced into the previously static view of organic nature. This led directly to the historical concept of man as the product of developmental process, and as the end product of a morphological process. This radically changed the whole conception of man from being outside nature looking in, to being a part of nature or the end product of a sequence of development. Man became part of the object of science.

The meeting of these three lines of thought: the organism–environment concept in biology; the subject-object interdependence in Kant, and the resultant concept of man as the object of science constitutes the man–environment paradigm of science and meta-science pervasive in nineteenth-century thinking and less so in the twentieth. We must not forget that the paradigm brings with it, through Newton's fallacious biological conceptions, the mechanism of Descartes in a revised form. This crystallises in the notion of the organism as a machine-system operating in an environment with which it is dynamically related by 'causal' connections. This is the concept generalised by von Bertalanffy and others as 'open system theory'. The 'open system' concept has little in it that is new, except a valuable emphasis on openness. It is basically the 'man–environment paradigm' raised to the level of an overriding explanatory principle.

The man–environment paradigm today and its paradoxes

Most ideas in science solve problems only to raise others. Often powerful ideas raise the most problems because they are applied widely. In many scientific fields today the man–environment paradigm is not so much a general theoretical idea, or even a meta-theory, but part of the mechanism of perception and understanding itself. It seems so natural that it has become invisible to us. Our argument is that the concept embodies profound anomalies. If we are right, then we would expect these difficulties to appear continuously in research as fundamental, unresolvable difficulties, even paradoxes. This would be made worse by the invisibility and 'naturalness' of the paradigm assumptions because we would look everywhere except at the most basic ideas for the source of difficulties and lack of progress. Precisely this situation prevails in many branches of social sciences and environmental studies. Social sciences appear to be progressing through a conflict of rival theories. These, although irreconcilable, are not in conflict. They represent the opposite sides of the paradoxes inherent in the paradigm from the beginning.

The basic argument is simple. Although the man–environment paradigm initially promises a reconciliation of rationalism and empiricism (that is, the basic problem of how far knowledge and behaviour originate within us or outside us), it only succeeds in reproducing it in exacerbated form. The two views are not reconciled, as are, for example, the notions of individual and population in modern genetic theories. They are merely locked together in a mechanical unity. Two mutually exclusive epistemological positions – that of the organism looking out into the environment, and that of the environment bearing in on the individual – are created. Each offers a radical critique of the other's inadequacy; each exists as an anomaly within the explanatory systems of the other. Individual, or creative, behaviour appears as an anomaly in environmental mechanism. Subjectivist views cannot begin to explain the coercive nature of, for example, social structures. Explanations within either scheme appear radically incomplete. Taken as a whole, the paradigm cannot account for individual or social knowledge, or for the nature of societies, or even urban systems, except by reproducing the mechanism and metaphysics of its predecessors.

The sciences of man are, with important exceptions, divided between those that take the position of the subject first and foremost, and those that take the position of the environment as constructive of the subject. In spite of their apparent conflict, both positions are the products of the paradoxes of the basic epistemological paradigm from which they developed. But the contest, while being unreal, is also unequal. The mechanistic notions gave a natural predominance to mechanistic interpretations of the environment which were part of the foundational apparatus of the paradigm. The paradigm thus appears in particular sciences as a fundamentally mechanist-empiricist programme harassed by subjectivist objections. This is most clearly illustrated in psychology where the dominant programme for 50 years was behaviourism, which in its extreme form seeks to explain all behaviour in terms of 'stimuli' from the 'environment' into the 'organism'. Throughout its history behaviourism has been complemented by a rival, quasi-scientific subjectivist school of psychology, which begins by asserting that behaviourism and 'science' generally can neither deal with the complexity and individuality of human behaviour, nor with such categories as free will. This line of philosophical psychology is chiefly associated with names like Merleau-Ponty, Sartre, and other figures dismissed by behaviourism as 'introspectionist'.

The pattern is repeated in sociology. Comte, the generally acknowledged founder of a specifically scientific sociology, worked on the basis of the organism–environment notion. His version is more clearly mechanical than any of the eighteenth-century models. From Comte's foundational work, developments in sociology are easily understood in terms of oscillations between the poles of paradox generated by the man–environment paradigm. On the one hand, writers like Weber attempt to construct 'society' from the socially significant actions of individuals, holding that the concept of society can have no meaning over and above the interactions of the aggregated individuals who are its members. This is the trend known as 'methodological individualism', which appears again and again in the subsequent development of sociology. On the other hand, writers like Durkheim take a contrary point of view, looking at the individual from the point of view of society. Social 'facts' can be recorded 'objectively' as in every other science. It is the structure of these facts which determines individual behaviour, if only on the basis of probability. Weber denies society. Durkheim reifies it by making it into a concrete 'thing'. Each point of view looked at intuitively appears equally absurd. How can a subjectivist point of view explain Detroit or Dagenham? Or how can an objectivist point of view explain that human beings do not have to obey the laws of the social structure, but are able to think and decide within limits on alternative futures? The fault is in the underlying paradigmatic notion of a division into a mutually exclusive world of subjects and objects, organisms and environments.

This basic dispute in the two positions in sociology, although laid down long ago, continues to reappear in new forms and is continually represented as though it has only just been discovered (which is further evidence of the existence of a basic paradigm). For example, the current principal movement in sociology is based on what is called 'phenomenology', which in its most extreme form becomes 'ethnomethodology'. The movement is opposed to what we call 'positivism'. Positivism is characterised by a research methodology based on statistics, which appears to ignore the 'meaning to the subject and to the actor' on which the phenomenologists lay great emphasis. In fact, the phenomenologists are embracing the philosophy they claim to reject since they do not realise that 'phenomenology' (roughly: remove preconceived ideas and we will see the[16] essences of things as they really are) was, and is, a super-empiricist viewpoint. But they are also repeating a debate, in new terminology, ongoing in sociology ever since its inception in scientific form. The current philosophical debate in sociology is characterised by each party pointing to defects of the other's point of view, yet unwittingly sharing similar assumptions.

The same type of situation exists in philosophy, where the paradigm reappears in the form of the problem of subjects and objects, knowers and known. We have already pointed to how ancient epistemological problems reappear in strengthened form in the subject–object paradigm in philosophy. It is also relevant to look briefly at general questions in the theory of science which appear in several contexts. For example, the problems of 'nature versus nurture' and 'environmental determinism' both tend to have the same fundamental form derived from a conceptualisation of the world in terms of an organism and its environment. The problem appears in theories of scientific knowledge in similar form.

The conceptual apparatus of the mechanistic man–environment paradigm is pervasive in environmental studies. The paradigm treats the field of environment as scientifically describable set of problems. Yet at an intuitive level every attempt at a scientific approach to the environment appears, to some extent, to contradict common sense and even to offer the possibility of damaging our understanding of environmental phenomena. The current scepticism of many designers about the uses of research in environmental design may be well founded. The same conceptual apparatus permeates design as well as research. Some of the more blatantly 'scientific' attempts to interpret environment have come from design. Think, for example, of the Corbusian concept of the perfectly specified and perfectly efficient spatial environment matching the perfectly efficient activities of a factory full of workmen making regular and repeated movements, or Chris Alexander's revival of the same theme of 'fit' in a semi-mathematical form in the 1960s. Think of the concept of the 'man–environment interface' in environmental research being investigated by the 'architectural psychologists'.[17] Think of the supposed problem of environmental or architectural determinism, also the opposition to it stating there is no causal or determining relation, not that the form of the question is wrong. Yet it is on these grounds that many sociologists dismiss the environment as being of no importance. All these problems are not, as they appear, simple indigenous products within environmental studies. They are the results of their take-over of the general paradigm complete with its paradoxes. We admit that there is a special problem in environmental studies, where the object of study is human beings in their physical environment. But we must draw a clear distinction between the problem and the paradigm. There is no reason why the assumptions of the man–environment mechanism should necessarily be taken over piecemeal as the conceptual apparatus of the subject, although this appears to be what has happened.

Piaget's solution

One powerful alternative is offered by Piaget. His epistemological and psychological programme had been aimed at a scientific solution to the age-old problem of rationalism and empiricism – the riddle of how far 'knowledge' originates in the subject and how far in its environment. Piaget's solution is *interactionist* and *constructivist*. Piaget's constructivism shows that all these basic problems in the human sciences – nature versus nurture, genetic versus environmental 'causes' and so on – are insoluble in the terms in which they were set up. They only appear as problems because of bad formulation resulting from the assumptions of the man–environment paradigm.

Piaget begins from the notion of organism–environment, but he adds to the primitive empiricist or behaviourist account in three important ways. First, he argues that 'knowledge of an object does not consist of having a static mental copy of the object but of effecting transformations and effecting some understanding of the mechanisms of these transformations'.[18] Second, logical relationships on which these transformations are based have their origins in physical operations, or more precisely in coordinations of operations such as uniting, ordering, introducing correspondences and so on. Third, knowledge is not therefore determined either by the knower or the known, but by exchanges and interactions between them. 'The fundamental relationship is not one of simple association but of assimilation and accommodation. The knower assimilates objects to the structures of his actions (or of his operations) and at the same time he accommodates these structures (by differentiating them) to the unforeseen aspects of the reality he encounters.'[19] Piaget thus places great emphasis on the subject's own constructive activity, whether he is performing on objects external to him or on internalised symbols and representations. He constructs a model of cognitive development which avoids the rationalist concept of 'innate ideas' and the empiricist alternative of determination by the environment. This is very convincing, and in its general lines, if not in all experimental details, it appears to be well supported by the experimental evidence amassed around this general theory.

As an account of individual development, it appears to be relatively complete. As a total account of epistemology, it is perhaps less so. It does not, for example, tell us how societies construct the category of knowledge, bearing in mind that these are the structures that the developing child actually learns. Taking language as an example we may reasonably question Piaget's use of the term 'constructive' – constructive of what? Language already exists. If it did not, the child would not learn it. In an important sense, what the child acquires by his 'constructive' and 'transformational' activities is a copy of a structure that already exists, even if it is imperfect in all its details. The child can use it creatively, and say things that are completely novel. But he can only do this because his cognitive structure of language is a copy of the structure of language as it exists outside him both in time and space. The problem of the 'copy', having been removed from the scene at the level of atomistic inputs into the organism, reappears at the level of the total structure. We believe that a consideration of this point can take us beyond Piaget's solution to the total problem of epistemology, and at the same time guide us towards a genuine escape from the paradoxes in the man–environment paradigm.

Langue, parole and non-spatial space

Both in its origins and its paradigmatic influence on science, the man–environment paradigm was a *spatial* concept. Its geographic and mechanistic components confirm this. Newton's 'ether' linked the parts of space together so they could work as a machine. A machine is essentially spatially unified. The 'environmental influence' of the geographical component of the concept was spatially specific. This was its whole point. Similarly in behaviourism, the environment of the organism is its immediate spatial environment. The fundamental anomalies introduced by its spatial assumption have been totally overlooked.

This assumption and its mechanistic implications themselves generate many mysteries that appear unresolvable within the paradigm. For example, it is said that society cannot be more than the sum of individuals who comprise it, since these are the only 'units of analysis' that can be seen. Society cannot exist since it cannot be seen as a unity in real space. It is therefore the aggregation of a large number of individual iterations which can be seen in real space. On the other hand, intuitively we know that society exists even if you cannot see or touch it, and that environment is somehow important even though it does not determine society in any mechanistic way. The point is that the fundamental scientific concepts which we use for analysis themselves construct these problems which do not exist at the level of common sense nor need they exist in science, as we hope to show.

A similar situation exists in environmental studies. Because non-trivial 'causal' relations between environment and behaviour cannot be found, physical environment is dismissed as of little importance by 'behavioural scientists'. At best this is old-fashioned, a repetition of the 'action at a distance' debate. At worst it is self-induced blindness, since only certain types of relations are allowed to be scientifically significant by the paradigmatic assumptions. It overlooks completely the importance of symbolic and semiological systems in the human universe, and the form of organisation of such systems which is not spatial, but logical.

This is a difficult distinction, but a necessary one. It is only possible to render architectural space accessible to science by first destroying the concept of space at the paradigmatic level. A nonspatial alternative must be constructed to the naive spatiality of the man–environment paradigm. The movement known as 'structuralism', in its scientific forms (but not in the form in which it has so far been presented in architecture) is founded on such a paradigmatic alternative.

It proposes a simple yet profound change: the paradigmatic substitution of logical space for spatial space. This sounds initially strange, but only because of the assumptions that underpin our thought. In fact, it is common sense. Logical space is an imaginary, many-dimensional space created by and filled with systems of signs, symbols and representations. It exists neither purely in our heads nor in real space outside but constitutes the medium through which the relation between the two is made. It works both ways. Logical space creates spatial or architectural space as one of a number of perceptual 'realities' it interprets. It represents the results of our cognitive operations on the world, including the world of artificial things. It is to be reconstructed scientifically through structural mathematics – principally algebra and logic – rather than quantitative mathematics – arithmetic, calculus, statistics – of the man–environment paradigm.

We exist not in 'spatial space' pure and simple, but in spatial space as it has been constructed in terms of the contents and structures of logical space. This has happened, Piaget-wise, through our cognitive activity by which we have made sense of the world, retaining as we go the structure of that understanding, and developing it to assimilate new experiences as they occur in real space. Artefacts exist as functional objects, but they also exist as representations, as signs and symbols. Due to our mode of existence by which representations are organised into structures, to make them intelligible, the variability of the environments which are presented to us in real space are understood by the structures which have been laid down in logical space. The relation of man to artificial environment is primarily a cognitive relation. It is mediated as are other cognitive relations by the organisation of representations in systems whose structures constitute the means by which new experience is made intelligible, but which are always open to further development and restructuring. The urban space we see is intelligible because it is part of a total structure that does not exist in spatially unified form. It, like language, can be thought of as existing in logical space. Individual understanding is a transformation of a structure that is invisible, like all structures, but no less real than the structure of language is real.

The same may be said of society. Society exists for us in terms of functional processes which involve us and ensure our physical survival, and systems of signs and representations, which are equally coercive because they constitute the only means we have of making society intelligible. Although the symbolic representations (by which we recognise the existence and coercive power of society) are systemic for us (we have learnt, through our mental activity, to construct them as systemic), this does not imply that it is therefore a matter of personal interpretation. Systems of symbolic representation are to some extent coercive (but not invariably so) for they are vital to our whole mode of existence, our self-recognition as members of the larger unit which sustains us. We can only move 'outside' society by means of another complete theory. This is why revolutionaries need theories. To be free of the symbolic representations requires us not only to restructure our perception and understanding but also to construct an alternative semantic for the symbolic systems of society – an alternative theory which will construct a different picture entirely of the social 'reality' represented in symbolic representations. Society does not emerge from the interaction of aggregations of individuals. Society is an artefact passed on from one generation to the next, changing and occasionally transforming structurally as it goes. It is embedded in orders of functional processes and symbolic systems which, in constituting the means by which social existence is rendered intelligible, constitute also the means by which we are controlled by it. It is not that society has artificial functional processes and symbolic systems. These *are* society. Individuals 'pass through' society, transforming it a little just as they 'pass through' the structure of language which was there before they came and which will remain, slightly modified, after they are gone. Much of life is lived at the representational and symbolic level. To make life and other things intelligible, representations and symbols must be organised into systems and structures. We can admit the firm existence of 'society' in the same way as we can admit the existence of language or environment. In sociology itself, it turns out that the implicit spatiality of the man–environment paradigm mechanics has made 'society' disappear.

Let us look again at language to clarify what we mean by the relation of the thinking individual to these abstract structures existing in logical space, but not in real space. We have said that the language which the child learns already exists. We may ask: where? Within the spatial assumptions of the man–environment paradigm, the answer must be: nowhere as a total structure. But this is evidently absurd, a paradigm result, not a scientific conclusion. Language exists as a structure – albeit a slowly changing one – embedded within the total set of systems of signs and representations which constitute our mode of being and mode of structuring our existence. These systems of representation also bear some, as yet poorly understood, relation to the patterns of functioning in which we survive and sustain ourselves in our accustomed modes of existence. The child in learning language does not equilibrate to a series of atomistic inputs but to this pre-existing structure in logical space. This manifests itself in perceptible form in all the fragments which the child hears. The child deduces a structure which already exists from the patterns of the fragments with which the child is presented. The learning of the structure enables him to deal more and more creatively with the fragments of sound that are heard. Very quickly indeed he will both interpret and create novel utterances.

If we take this as our model for environmental cognition, then many of the problems of the man–environment mechanism disappear. There is no problem of causality: no problem of determinism, no problem of uniqueness and sameness, both of which exist in language; no problem of subjects and objects, since the paradigm states that the individual has learnt in use a structure which already exists and which he may within limits transform according to his wishes and ability.

From a scientific point of view the paradigm itself has changed. We are looking for different kinds of order, dispensing both with mechanistic 'man–environment systems' and totally autonomous subjects, voluntarily assembling into larger unities. Instead, we find that we are assuming a paradigmatic distinction of a different kind: between 'langue' or 'language' (that is, the total structure of a language as it exists at any point in time) and 'parole' or 'speech' (the use individuals make of that structure, modified by their codes).[20] In this paradigm, the structures of the 'logical environment' and those of the 'individual' are related in a fundamentally different way by the relation of transformation on some deep structural identity. We approach the individual through the structures in logical space which means his existence, not his 'subjectivity' or his 'behaviour'. Behaviour will be what it should be in science: the evidence for or against theoretical constructs.

Translated into environmental studies the 'langue-parole' distinction becomes the morphology of physical environment (inevitably, a social account of the rules governing the production and transformation of space) and the coded transformation as it appears in individual behaviour and cognition. The problem field generated is one of deep and surface structures, variable codes and transformations which are better able to come to grips with environmental reality and experience than either man–environment mechanism or its subjectivist alternative. The latter gives a problem field of stimuli and responses; the determination of causal relations between environment and man; the measurement of crude spatial influences and the search for relations at an alleged interface. Eventually and reluctantly, 'meaning' is allowed to exist as an 'intervening variable' between stimulus and response, but converting it thereby to an a priori subjectivism into which the man–environment paradigm always falls when the empiricist model of determinism is abandoned as the principal strategy. This, as we have argued, is merely the other side of the paradox that the paradigm itself creates, and must not be thought of as an advance. We merely move from determinism into relativism, and all hope of finding structure in the situation is lost.

What we are saying comes down to this. People, in some sense or other, and however variably, understand the environment, or understand something through the environment, or are significantly baffled because they do not so understand it. When they say things about the environment, they actually mean something and we ought to set about finding out what it is. At an intuitive level, people continually exchange ideas about the environment, which they appear to find meaningful. For some reason, possibly because surface meanings remain indeterminate, these formulations are rejected by environmental science, when it seems to us they should constitute one of its theoretical starting points and one of its chief empirical objects.

We are not simply talking about 'participation planning' or forms of 'action' research or taking verbal performance at its face value. We are talking about a structural approach to environmental understanding, as applied to 'primitive' societies by such as Lévi-Strauss and Leach, and pioneered in terms of our society by Bernstein. Structuralism, whether in the study of languages, society or environment, is the recapture of the logical structures which make 'meaning' possible, the conscious models which generate the richness and diversity of man-made artefacts and also make them intelligible to us. It is the reconstruction of what we already know. It is the theory of everyday life. It is no less the 'machines for the production of variety', foreshadowed by Heraclitus and Leibniz long before such an idea could be formulated logically or algebraically.

The astonishing thing is that it is not possible to formulate what we have just said within the scientific terms permissible in environmental studies, dominated as it is by the techniques and paradigmatic assumptions of the 'related' disciplines. All we have tried to do here is to find a scientific basis for saying the obvious. The length of the journey we have had to make underlines the fundamental nature of the paradigmatic difficulties which the discipline has struggled with so unsuccessfully for so long.

About the authors

Bill Hillier and Adrian Leaman are members of the RIBA Intelligence Unit and teach an MSc course in environmental theory at University College London. The authors would like to thank Roy Oliver for his comments and suggestions on the text.

Notes

1. F. J, Odling-Smee, K. N. Laland, and M. W. Feldman, *Niche construction: The neglected process in evolution* (Princeton, NJ: Princeton University Press, 2003).
2. M. Heidegger, *Being and time*, translated by John Macquarrie (Oxford: Blackwell, 1927) (1962 edition).
3. R. Llewelyn-Davies, 'Deeper knowledge: Better design'. Paper given at Conference on Architectural Education, Magdalen College, Oxford, April 1958.
4. M. Foucault, *The order of things: An archaeology of the human sciences* (New York: Pantheon Books, 1970).

5. T. S. Kuhn, *The structure of scientific revolutions* (Chicago: University of Chicago Press, 1962).
6. A. Ravetz, *Model estate: Planned housing at Quarry Hill* (Leeds: Croom Helm, in association with the Joseph Rowntree Memorial Trust, 1974).
7. J. Piaget, 'Developmental psychology', in *International encyclopedia of the social sciences*, edited by D. L. Sills (New York: Macmillan, 1968), pp. 136–147; Piaget, J., *Biology and knowledge: An essay on the relations between organic regulations and cognitive processes*, translated from the French by Beatrix Walsh (Edinburgh: Edinburgh University Press, 1971), pp. 136–147.
8. F. de Saussure, *Course in general linguistics* (New York: McGraw Hill, 1959) (1966 edition).
9. It is perhaps notable that no reference is made to Karl Popper's notion of 'three worlds', which seem directly relevant to this latter discussion, bearing as they do on the real, mental and cultural worlds of knowledge (K. Popper, 'Epistemology without a knowing subject', *Studies in logic and the foundations of mathematics*, edited by B. Van Rootselaar and J. F. Staal (Amsterdam: North-Holland Publishing, 1968), pp. 333–373. Editors' note: This essay is reprinted as chapter 3 of Popper, *Objective Knowledge* (Oxford: Clarendon Press, 1972), a volume that Hillier owned. We know that Hillier was reading Popper as this had formed a central plank of the epistemology he outlined in 'Knowledge and design' to the Environmental Design Research Association conference in Los Angeles the previous year (B. Hillier, J. Musgrove and P. O'Sullivan, 'Knowledge and design', in *AIA-Architect-Researchers' Conference. Environmental design: Research and practice. Proceedings of the Third Conference of the Environmental Design Research Association*, edited by W. Mitchell (Los Angeles: University of California Press, 1972), pp. 29-3-1 to 29-3-14).
10. B. Hillier, A. Penn, J. Hanson, T. Grajewski and J. Xu, 'Natural movement: Or configuration and attraction in urban pedestrian movement', *Environment and Planning B: Planning and Design* 20 (1993): 29–66. 'Natural movement' is still the most cited journal article in the field.
11. R. Willmott, 'Structure, culture and agency: Rejecting the current orthodoxy of organisation theory', *Journal for the Theory of Social Behaviour* 27 (1997): 93–123. Willmott's critique was levelled at theories such as Giddens' 'Structuration' and Bourdieu's 'Habitus', which he says aim to overcome possible dualisms by conflating the strata of society.
12. G. Canguilhem, *La connaissance de la vie* (Paris: Vrin, 1965), pp. 129–154.
13. *Newton's philosophy of nature: Selections from his writings*, edited by H. S. Thayer and J. H. Randall (New York: Hafner Publishing Company, 1953), p. 54.
14. See *Newton's philosophy of nature*, pp. 135–180 (eds).
15. G. Buffon, *Histoire naturelle*, vol. 1 (Brussels: Lejeune, 1749).
16. Correction of error in the original.
17. Correction of error in the original.
18. Piaget, 'Developmental psychology'. Quotation from 'II. A Theory of Development', p. 140.
19. J. Piaget, *Biology and knowledge* (Edinburgh: Edinburgh University Press, 1971).
20. Saussure, *Course in general linguistics*.

3 How is design possible? (1974)

The science of meaning and the meaning of design. Introduction to 'How is design possible?'

Sean Hanna

In this paper, which predates the first articulation of space syntax by two years, Hillier and Leaman set out a founding idea of the field: that the central question of how design is possible is a scientific one, and that it can be posed as such by framing it in the context of linguistics. They draw explicitly on Chomsky's linguistic model,[1] which proposed that all languages share a common structure, and it is this that allows the particular rules of any individual tongue to be acquired. This casts linguistics itself as a scientific domain by isolating this structure for examination, in part by declaring syntax independent from semantics. The syntactically well-formed sentence, 'Colourless green ideas sleep furiously', Chomsky claimed, is at the same time semantically nonsense.

The empirical task for Hillier and Leaman, as it was for Chomsky, is then to describe the equivalent deep structure of the environment. By considering structure itself as the object of study, the question of its performance, the processes by which an actual designer (or speaker) acts, could be ignored in favour of modelling only idealised competence, the fit of the environment (or utterance) to that structure. The paper begins with a call to study the environment itself instead of how designers design and then immediately sets out to define those basic units of environment: the pre-structures that underlie our experience of spaces and walls, their transformation into built form as a biological analogy of genotypes specifying phenotypes, and the links between our interpretation of the world to the one we construct, which Hillier and Leaman term 'manifolds'. Space syntax still remains a domain of competence models – observations of aggregate movement are associated directly with properties of space, without requiring a psychological explanation.

But Hillier and Leaman's departure from Chomsky is crucial – their elementary units are complex and connected. They insist these are not atomic tokens, such as representing words in a linguistic grammar (Chomsky's syntactic structure was introduced precisely to allow the combination of word-atoms), but ascribe to them a rich complex of social relations, a network of connections they each carry and instantiate, as a whole, each time they occur. Meaning is intrinsic to them, as it would be in Hillier's later morphic languages, but instead of the arbitrary association of sign to referent in spoken language, meaning is the natural structure of these social relations. This structure should become the focus of the new science.

In proposing this, the paper is a critique of then-contemporary notions of design (CAD technology of the 1970s, the design methods movement), and also offers a view toward future developments and an argument that is just as relevant today. The activation of a complex manifold structure as a whole is much more closely aligned with today's artificial neural networks than with symbolic AI prevalent at the time. Decades later, models of language production would be developed that acknowledge this complexity to provide more neurologically plausible mechanisms than a purely syntactic grammar. Implementations like Elman's recurrent neural network[2] produce human-like results, rely on evolution over time and suggest the distinction between syntax and semantics may be illusory – the two are simply differences of degree, and each use of a word necessarily carries a whole structure of meaning as well. It is the AI descendants of these that today generate plausible, human-like text.

Yet in design, the paper anticipates a critique of twenty-first-century CAD and building information modelling (BIM). This adds additional layers of data to objects, which aspire to the deeper manifold structure proposed, but they are still distinct units defined by phenomenal, superficial geometry. Hillier and Leaman stressed the importance of the evolution of structures over time, while BIM has become ever more committed to standards and Industry Foundation Classes (IFCs) that are not meant to change. Their prescience may explain the limitations of BIM in serving architects' needs, particularly in early, creative stages of design.

The paper concludes with two questions. The first asks 'how does the environment influence social behaviour?' and the second, 'what social behaviours determine the environment?'. Hillier and Leaman criticise the first as merely a manager's question, but it is worth considering the degree to which the quotidian practice of space syntax, in predicting movement and behaviour as a function of space, has remained within its sphere. The same is true of CAD and BIM. It is the second that constitutes the deeper scientific question and suggests that the structure of the complex of social behaviours embodied in the environment is still ripe for exploration.

'How is design possible?' is a leap towards the scientific framing of linguistics which would become space syntax. It originated specific ideas which would later develop into core concepts and sets the basis for subsequent analytical methods: the production of p-model from g-model would become an inverted genotype and manifolds and pre-structures would become morphic languages. In retrospect we see in the paper some concepts as green shoots that have since flowered into mature space syntax theory and practice, while other questions remain relevant and even more pressing now coloured by five decades of technology. If in 1974 this potential was yet to awake, reading now we see these formerly colourless green ideas sleep furiously.

How is design possible?

Bill Hillier and Adrian Leaman

Originally published: Hillier, B. and Leaman, A. 1974. 'How is design possible?' *Journal of Architectural Research*, 3, 4–11.

Introduction

It has been observed that the art of mathematical proof lies in finding a framework in which what one wants to say becomes nearly obvious. Improbable as it sounds, the search for theories in science is often similar. What has to be 'said' is the observable evidence and its manifest relationships. A theory is an abstract model which predicts that such observables and such relations will occur. A 'good' theory renders data 'obvious' which a 'bad' theory renders contradictory or disconnected.

The theories of design developed over the past decade in connection with 'systematic design', 'design method' and subsequently 'computer-aided design' do not in general have the merit of rendering the evidence about design 'nearly obvious'. On the contrary, they make it appear mysterious. For example, the syncretic generation of outline solutions in the earliest stages of design is made to appear illegitimate and undesirable on the grounds that any 'rational' approach to design must seek to generate the solution as far as possible from an analysis and synthesis of problem information and constraints. The penalty of not following this 'rational' procedure is that bad old imagistic methods of design will be perpetuated. 'Rationality' in design was virtually equated with purging the mind of preconceptions, to make way for a problem-solving method which linked a procedure to a field of information.

The sources of this conception of rational design were analysed in a previous paper, in which it was also argued that the designer's 'pre-structures' were not at all an undesirable epiphenomenon, but the very basis of design.[3] Without pre-structures of a fairly comprehensive order, it is not possible to identify the existence of a problem, let alone solve it. The task now is to find a scientific framework in which such a formulation would become 'nearly obvious', and this will be equivalent to a theory of how design is possible. In general, it will be argued that design is possible for roughly the same reasons as *speech* is possible. Methodologists have not been able to characterise speech because they have tried to look at it independently of a much larger structure analogous to *language*. Too small a system has been considered.

Genotypes and phenotypes

Consider an example. An army marches all day. At nightfall a halt is called and unpacking begins. Within a short time, a structured environment appears. Tents of various kinds and sizes are placed in certain definite relations; kitchens, sentry posts, flags, fences and other paraphernalia are erected. A complete environment is, as it were, 'unfolded'.

The army experiences this as a simple, repetitive procedure. But looked at scientifically in terms of the structures which must exist for such a simple 'unfolding' to be possible, it is both complex and illuminating. First, the observable environment which is unfolded is based on a set of instructions in the army manual or in standing orders. Second, these abstract instructions and relations are 'embedded' in the items the army carries about with it to manufacture this environment, which may be called its 'instrumental set'. Third, and most important, it will be noted that any series of camps unfolded by the army will exhibit both similarities and differences. In this simplified and artificially deterministic example, the source of similarity is the set of instructions as embedded in the instrumental set, and the source of differences is the local constraints and contingency, which will include personal and environmental factors, strategic considerations and so on.

This example has the curious attribute of resembling a theory in reverse. Just as scientists try to make abstract invariant models which 'generate' and thereby explain a variety of observable phenomena, so in this case, the abstract model is the starting point and the observable phenomena are generated. One can speculate on a Martian anthropologist's theories as to how the similarities and differences of series of such camps are to be explained. One can imagine, for example, the difficulties that a purely 'environmentalist' theory would encounter in explaining the observable phenomena.

Although an extreme example, this structure does illuminate certain pervasive properties of cultural forms. In many situations, sets of basic instructions which are variably unfolded are not written down nor genetically programmed, but embedded in the artificial systems which we call 'cultures'. Deep cultural structures may be transmitted unchanged through several generations, yet produce great variety at the observable level. Such underlying stable structures correspond to what biologists call 'genotypes', as compared to a 'phenotype', which is a variably developed observable form. In the army example, the genotype is the information carried in the instructions and embedded in the instrumental set; the phenotype is the observed layout and activity of the camp. In most cultural areas the structured sets of instructions, the genotypes, change and evolve much faster than biological genotypes, and may even be able to alter themselves by self-reference, but the overall *logical* form is the same. In order to avoid confusion with the more closed and stable genotypes of biology, these looser cultural genotypes and phenotypes may be referred to as *g-models* and *p-models*.

An important property of g-models is that they become unconscious or *autonomic*. Moreover, they operate not only as a basis for creative action but also as a basis for understanding and interpretation. They act, literally, as theories of the artificial systems they interpret. In all human societies, such theory-like structures are evolved to mediate the artificial, natural and social universes. Language, for example, uses an unconscious, autonomic, syntactic and semantic g-model as the basis for speech. It is this that makes Chomsky's 'rule-governed creativity' possible. In general, the richness and variety of intelligible cultural existence is founded on the existence of such models. But it is only as scientists (or artists like Marcel Duchamp) that we attempt to bring them into the light of day.

Such theory-like models are no less necessary in design. But the fact that they are autonomic may lead the designer into a curious illusion when he tries to explain the nature of his own activity. It might appear to him that the phenotypical choices he makes in relation to his analysis of the constraints of a particular design problem could, or should, account for the *entire* solution, that is for its similarities to the g-model as well as its p-model differences and variability. This is exactly the illusion conjured up by 'scientific' approaches to design. As would be predicted, it is the source of the solution itself which cannot be described in terms of the methodologies, and it therefore appears extra-rational. Because the g-model is both the structure by which the designer analyses and identifies as well as solves his problem, it has become invisible to him. Nevertheless, in any analysis of how design is formally possible, such structures are of the utmost importance. To ignore them in representing design as a decision procedure is like assuming that a speaker re-invents semantic and syntactic structures which he depends on knowing in advance in order to use and understand the language. Most of the fallacies of behaviourism, as well as design methodology, can be derived from this extension of the domain of scientific interest.

A simple general theoretical framework may therefore be suggested within which the everyday evidence about design becomes 'nearly obvious': design is the search for the appropriate transformation or unfolding of pre-structures (however rapidly these may be evolving) in relation to the constraints imposed by the environment of the problem. *Design is therefore both the transmission* (g-models) *and transformation* (p-models) *of pre-structures*, a process of elaboration and discovery, within which every solution may be unique. Referring obliquely to H. A. Simon, it may be argued that design is simple, only its environment is complex.[4]

The pre-structures on which design depends are synchronous – they are not arranged in a time sequence. But design obviously occupies time and contains certain sequences of events. The decision procedure approach to design method ignored the importance of the synchronous plane in structuring the design process. On the other hand, a pre-structure that is not subject to operations cannot be transformed into a unique p-model. The problematic of design method studies is therefore twofold: to characterise the autonomic pre-structures by which the designer interprets his problem and which also act as a 'solution field'; and to characterise the operations which may be performed with, and upon such, structures in a more or less complex problem environment to produce unique and effective solutions. The first is a problem of synchronic analysis, the second, diachronic.

This analysis will suggest that design is a relatively simple set of operations carried out on highly complex structures which are themselves simplified by 'theories' and modes of representation. It suggests strongly that if design methods are to be improved, then it is more important to study the environment itself than how designers design, since in the last analysis the designer's understanding and how the total g-model constructs his understanding is the most important factor influencing design. Insofar as design should be studied, it should be in terms of how the designers' internal models transform environmental reality, and what conventional simplifications and distortions are introduced by the use of methods of representation and selection of problem variables.

The structure of a solution field

Understanding the structure of solution fields, or at least their organising principles, is therefore a major problem in design theory. If no models yet exist, comfort may be taken from the great generality of this type of problem. Most man-made systems on which our daily lives depend combine the paradoxical properties of being 'known' in the sense that we are able to use them without mistakes to convey inter-subjective 'meaning', but being at the same time inaccessible to scientific explanation. Languages, cities, fashions, economies, societies and even theories all fall into this category. We have 'sciences of the artificial' to enable us to 'understand' what we already 'know'. In all such cases, it is likely that intelligibility at the everyday level is due to autonomic underlying models; lack of intelligibility at the scientific level is due to our not having adequate scientific models for such models. Recent research in linguistics, artificial intelligence and mathematics suggests lines which research might follow.[5] It may at least be shown that certain types of structure must be inherent in such fields in order for them to work in the way they do.

It is useful to make an initial subdivision of the problem into three. First, understanding *how the solution field is organised formally*; second, understanding *how it is organised semantically*, that is, understanding the 'theory' or 'bundles' of theories and g-models on which it is based; third, understanding how it is *represented*, and how these representations may be acted on so as to produce effective transformations. In certain senses, it will be seen that all three are the same problem.

Most artificial systems exhibit a duality between a morphology based on entities which are manipulated – words, walls, coins, mathematical symbols – with a formal or logical structure which relates such entities to other entities of the same type, and also to other dissimilar entities like 'meanings' or 'values', by which social signification is achieved. In general, formal structures, like grammars, for example, are autonomic while morphologies make up a consciously manipulable set.

The key to understanding how such structures combine into effective artificial systems must tell us about the nature of the morphological 'units', the building rules of the overall logic, and how the two combine together in a 'natural' way. The thesis here is that the elementary units of the morphology are not 'units' at all in the usual sense, but are already structures. Moreover, their structure, as opposed to their phenomenal form, is as autonomic as the overall formalism. The understanding of all such systems lies in discovering how the internal autonomic structure of the 'simplest structures' of the morphology already contains the rules which govern aggregation into higher logical forms. The failure of general system theory to progress beyond an elementary level in characterising how such systems work is because this elementary principle of the dynamics of artificial systems cannot be formulated within a definition of a system as 'elements and their relations'. There simply *are no elements*.

This abstract argument may be made intuitively accessible by considering a simple example: that of a coin. Consider what it is that makes a coin more than a piece of metal. A coin depends on entities and formal structures which are both absent physically, and which may be entirely abstract. The 'meaning' of a coin depends on the existence of logically *similar* entities with which the coin may be compared (coins of other values); a set of *dissimilar* entities to which it may be systematically related (loaves of bread, a day's work); and a set of rules for making transitions between and within these two domains. The coin 'carries' with it at least so much structure which gives this 'simplest structure' the form of an H, with the left vertical representing the similar domain, the right vertical the dissimilar domain, and the bar the mapping structure (Figure 3.1).[6] The H shape is given for reasons of intuitive accessibility, but from a formal point of view, the simplest structure should be represented as a commutative square (Figure 3.2), since transformations in one domain carry with them transformations in the other if 'meaning' is to be retained. Such commutative squares, which appear to be general simplest structures in artificial systems, are fundamental to modern algebra and category theory, and offer a simple formal representation of the concept of 'meaning' (that is, a transformation in both domains such that the mappings are commutative).

This example provides an illustration of the general dependence of 'parts' on the 'whole' in such systems. In a sense, the internal structure of the 'simplest structure' is the same as the structure of the whole system. In architecture, the 'simplest structures' are even more complex, but have the same general form.

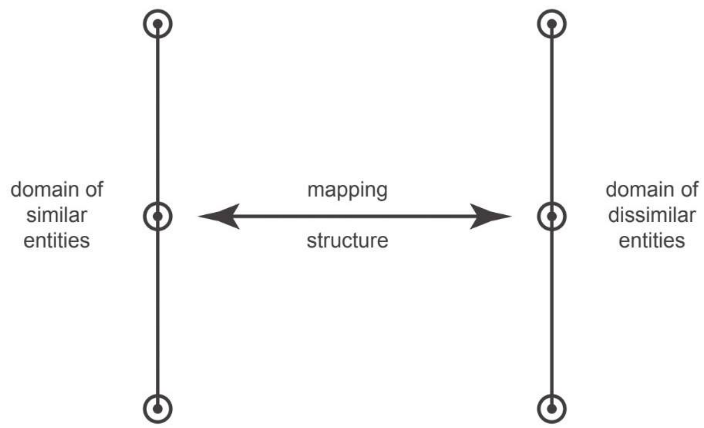


Figure 3.1 Illustration of 'simplest structure'.

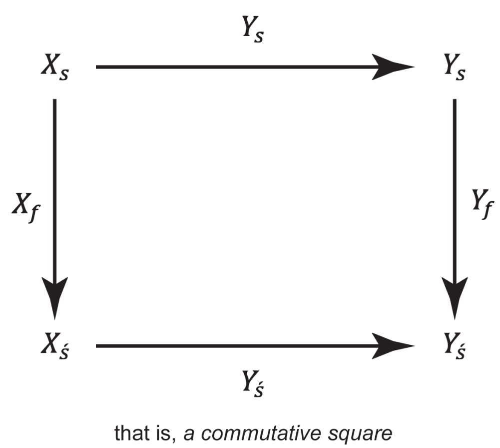


Figure 3.2 Simplest structure: algebraic form.

Consider, for example, a wall. A wall appears a simple object, but considered logically it turns out to be complex, even more complex, in fact, than the coin. In the first place, a wall, by reason of its purposeful existence, distinguishes at least two types of space. The reason for building the wall is to draw such a distinction. The same applies to the invisible 'walls' which are pervasive in artefactual forms.

Moreover, if the wall is perforated or not, then it also defines a relation between the two types of space. This is the simplest *phenomenal* form equivalent to the coin. To this must be added the relation to a similar domain of such structures, and dissimilar domains of usage and semantics to give the 'simplest structure' of architectural form. This may also be called a 'minimum meaning unit' (Figures 3.3 and 3.4).

The pattern of dissimilar domains of entities, with mapping structures between them, one domain of which constitutes the normal 'manipulable set',

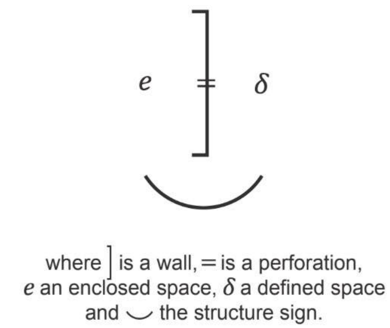


Figure 3.3 Architecture: simplest phenomenal form.

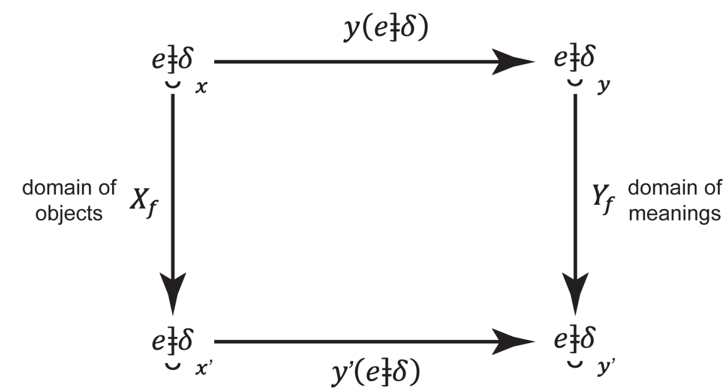


Figure 3.4 Architecture: simplest structure.

is very general indeed in human cultures. It might not be fanciful to define man as a structure who knows how to effect a large number of mappings between dissimilar domains which may be expressed as commutative squares. Languages have this form, with mappings between an abstract domain of 'meaning' and a concrete domain of phonetic sounds. The architecture of architecture is equally based on such structures which include, for example, the mapping between human behaviour and its spatial containment, or between psychophysiology and the environmental filter. In design, the mapping structures are *used* as autonomic devices to solve problems. In research, these mapping devices are studied in order to understand and improve them.

These arguments show, in principle, why it is possible for action on simple 'manipulable sets' to carry with it a high degree of morphological and logical complexity, often without our being aware of it. The encapsulation of the complex in the simple seeming object provides an initial tool for understanding the structure of a solution field. But this is already insufficient for design, since design is not carried out on objects but on representations of objects, whether on paper or in the head. It is necessary to understand how, formally speaking, this complicates matters, but again the complexity provides a key to how the system works.

It has been known since the Stoics that the ‘sign’ is only a sign by virtue of being connected to two dissimilar types of entity: the ‘lekton’ or conceptual referent;[7] and the real or phenomenal referent. Figure 3.5 represents these relations. This ‘simplest structure’ may be married to those already described to produce illuminating results. For example, ‘signs’ in the normal sense constitute a manipulable set. Insofar as they form a system in themselves, they do so by virtue of being based on some semantic model, or bundle of theories, about reality. There is of course no universal language in which every property of the world as it is can be equally and ‘objectively’ represented. Language is more like a theory of reality than a copy of it, as philosophers have argued. Languages act selectively on the world and at the same time systemically, and this is as true of mathematics as of natural language. In order to inaugurate the modern era in mathematical physics, Galileo purified the world of all properties that could not be represented in the evolving mathematical language, calling them ‘secondary’ because they were concerned with the perception of the subject rather than the measurable properties of the thing. This is still generally true of science. In classical physics, for example, the operations which may be performed on the calculus serve as models within sign systems for external realities, and thereby constitute the means by which phenomena may be said to be ‘understood’ scientifically. The recent advent of new mathematical formalisms like catastrophe theory[8] demonstrates how closely related the understanding of the universe is with the forms implied in mathematical systems. Modern theoretical physics accepts a basic duality in its research: on the one hand, it investigates formalisms; on the other, it investigates the physical universe itself, making models in the former which satisfy the appearances of the latter. The movement known as structuralism proposes an exactly similar strategy in the ‘sciences of the artificial’.[9]

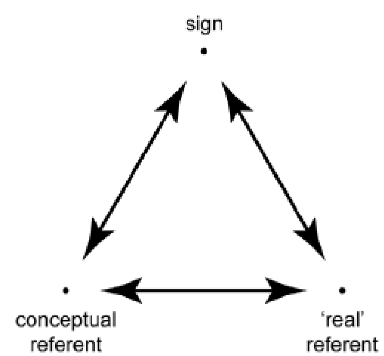


Figure 3.5 Sign: simplest structure.

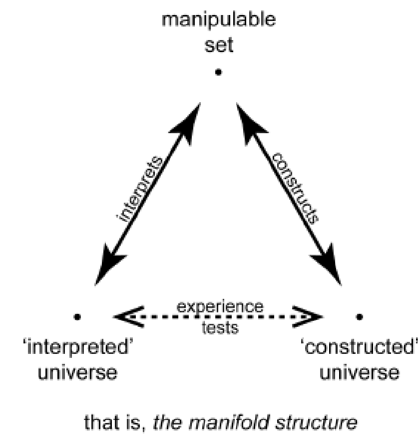


Figure 3.6 Sign system: simplest structure.

All sign systems have this duality. At the same time as *constructing* a permissible universe out of the maze of undifferentiated phenomena, they also systematically *interpret* it. This pervasive form of connectivity in the artificial universe is often overlooked. A term is needed and the notion of *manifold* may serve the purpose (Figure 3.6). Manifold structures are very general indeed. Consider a machine, for example. A machine constructs a permissible universe from which it can accept certain types of input. It then carries out a structured conversion on the input, and outputs an interpretation, which we call a product. An algorithm is similar. The algorithm constructs its permissible universe, performs a structured conversion on what is ‘selected’ within the domain so constructed, and outputs an ‘interpretation’. Like the previous examples discussed, the manifold form of connectivity (manifoldness) is one of the structures which makes everyday life possible. The form of connectivity of the manifold is such that when part of the manifold is activated, the whole necessarily also is.

The manifold is also one of the structures which makes design possible. Because the systems of representations on which the designer operates both construct and interpret their universe, and because they integrate the type of connectivity discussed previously, it is hardly surprising that design is largely a syncretic and hypothetico-deductive activity. The idea of the manifold in particular suggests that current approaches to computer-aided design may be involved in a kind of paradox. Even to name an architectural problem – say, ‘design a school’ – implies a whole range of solutions which will be more or less immediately activated in some sense through the designer’s manifold. The same applies to the definition of space at all levels. To try to relate abstract space to abstract activity overlooks the fact that each has already become an expression of the other. Concepts of space are already social to us, not pure perceptions waiting to be socialised. The designer, whether he is using the language of the brief, the language of concepts, or the language of drawings, is already embedded in a richly connected universe whose connections are those dissimilar domains that must be related in his design – these include activity and space, physiology and climate, and so on. Moreover, these structures are embedded in all the language the designer uses, and in the instrumental set – the technologies in terms of kits of parts and typical design solutions – to which his systems of representations refer.

Manifolds show why pre-structures must dominate design. But there is a further question. What is the relation of the manifold to 'reality'? With a manifold base, design takes an existing technological and environmental situation (kits of parts, solution types) as a necessary starting point. But this does not mean that design somehow works on the environment 'as it really is'. On the contrary, the distance between the 'real' socio-spatial situation and the designer's pre-structure which represents it will be subject to at least two types of transformation. The first will be concerned with the designer's 'code' – the structure which he abstracts from and interprets the given situation; personal history, professional education, the office, will all play a part in constructing this code. Second, a further transformation will occur through the influence of the method of representation itself which will act constructively and interpretatively as any manifold will. This means that the study of design codes and the effects of all methods of representation used in design are essential dimensions of the science of design.

These types of structure tell us in general how design is possible and what should be studied. On the whole, they are properties of all similar manifold systems of the artificial. But at least one other formal structure must be taken into account, which begins to distinguish the properties of architectural design from the general properties of manifolds. In all manifolds, the question of how function turns into structure is of central importance. In architectural design, this problem may be approached through the four-function model which, although presented elsewhere otherwise, turns out to be a structure in itself.[10]

The four-function model may be interpreted in the following way. Buildings mediate two different kinds of relationship between *man and nature* and between *man and man*. In mediating relations between himself and nature, man builds a climate modifier. In mediating relations between himself and other men, man builds a behaviour modifier. These two functions constitute the observable object-level realisation of the man-nature/man-man distinction. Each of these is mapped into a higher-order signifying system or, in other words, a general social language and accounting framework for such modifications. On the one hand, the object-level functions are mapped into a general analysis of resources whose language is economics, giving the resource function; this is a generalisation of the man-nature relationship. On the other hand, they are mapped into a general language of social signification, giving the symbolic function; this is a generalisation of the man-man relationship. Both higher-order languages comprise both lower-order systems. This structure is shown in Figure 3.7.

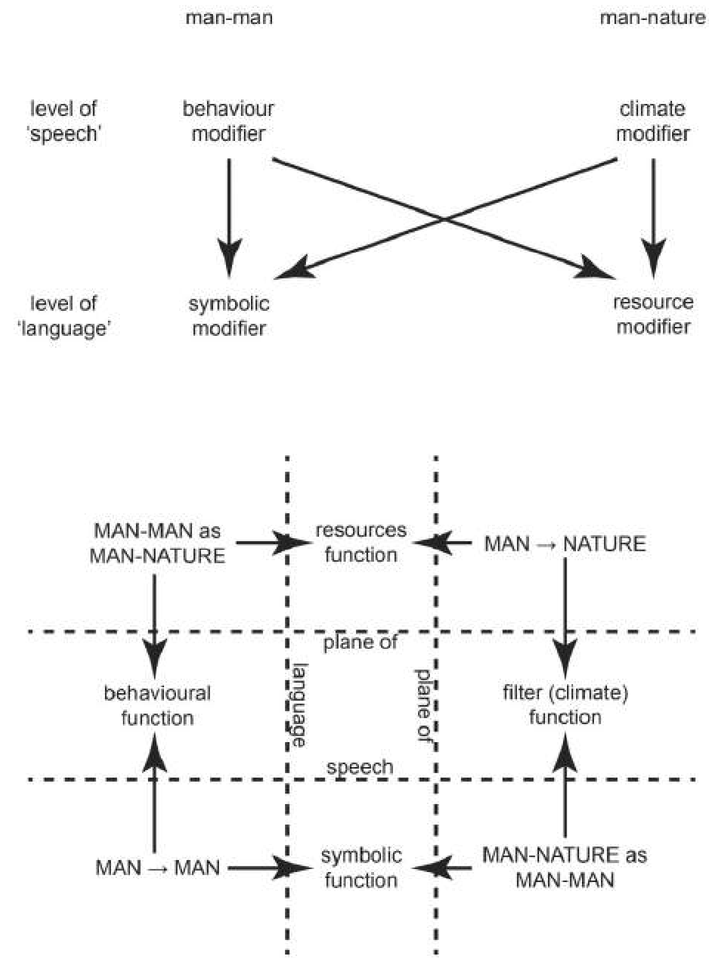


Figure 3.7 The four-function model as a structure (two versions).

Local and evolutionary time

Attention has so far been confined to the synchronous organisation of the solution field as a searchable structure and as a formal structure of relationships based on a general theory of environmental reality, through which problems may be both defined and solved, the same structure governing both definition and solution. As was previously argued, design may be seen as both the transmission of the abstract 'genotype' structure as a whole, and its transformation to realise a particular 'phenotype'. To understand how transformation occurs, the operations which are performable on the genotype structure must be understood. This, in effect, is equivalent to asking what time structures are relevant to an understanding of how design is possible.

Time is vital to an understanding of design at two levels. The first may be called the level of *evolutionary time*, where the cultural genotypes of architectural and urban form mutate gradually or are suddenly altered. This normally appears in the history book under the heading of the history of style, but it is clear that, from the point of view of the built environment as a whole, much more profound mutations in the relations of changes in society to changes in its spatial form occur. Evolutionary time is concerned with the gradual unfolding of a socio-spatial morphology, which includes the development and stabilisation of technologies, social processes of environmental control and all constructions of the mapping from social into spatial form. The second level is that of *local time*, in which the genotype is elaborated into a phenotype appropriate for a particular set of local conditions, through the activity of designers, users and others.

Design has usually been represented as a decision process in local time. The concepts of pre-structure, solution field and so on suggest that this is too small a part of the story to render the whole process intelligible. But since design does involve sequences of operations, in which time order is critical, then these must be elaborated. First, a general consideration. Time structures in design must be considered in relation to a ubiquitous property of manifolds – when part of the manifold is activated the whole also is. Consider, for example, in carrying out a conversation, the whole language structure is in action, both in order to interpret what others say, as well as to plan one's own.[11] The demonstrable dependence of the part on the whole requires this formulation. No part of the manifold may be used without having the whole at the ready. The manipulable set cannot be used without holding the rules as an autonomic generative structure.

Four types of local time structure may be distinguished in design. The first three have been dealt with at length elsewhere,[12] so the fourth – arguably the most important – will be the subject of attention here. First, a process of increasing specification, in, for example, a sequence of more precise drawings, continuously increases the range of excluded solutions and converges on a unique solution which will be built. This is the process of *variety reduction*. Secondly, the process of *conjecture-test* provides a basic unit of design activity by which activated parts of the manifold are simulated as solutions and tested against the problem constraints which will influence the form of the phenotype. Third, a time structure which is the product of the *complex of social relations* in which design is embedded. This leads to the sequential production of communications of various kinds, sequences of consultation and so on, all of which have to be distinguished from the underlying cognitive processes in the designer's mind, however much they may influence the result.

The fourth type of time structure is a *generalisation of the conjecture-test molecule* of cognitive activity to the overall level of the process in local time. It is here that the language analogy may be made most precise, and the language-speech distinction used to clarify what happens in design. What a speaker can say at any point in time depends initially on his competence, which is described by his grasp of the syntactic, semantic and functional structural abstractions which make up the language as a whole. He requires a systematic structure for mapping between meaning and verbal production. The limitations on this intervening structure will limit both what can occur in the domain of meaning and in the domain of sounds. The speaker, like the designer, starts from a pre-structure in which the most important entity is the abstract structure by which mapping between dissimilar domains may be affected. In all such cases there is a manipulable set. In the case of language, the manipulable set is words and phrases; in the case of the designer, it is the representations of the instrumental set. Without the mapping structures, these are virtually useless and only meaningless sounds or arbitrary artefacts could be generated.

What happens between meaning and speech happens in two stages, not one. First, a set of semantic units, known in principle to be sayable, is conjectured in the speaker's mind. This may be thought of as a transition from the domain of structured, combinable meanings, through the semantic structure, into a general set of proposals for speech which are as yet unspoken. These units are realised in the form of speech by passing them through a second mapping structure which constructs word orders, phrase forms, sentence forms and so on, out of conjectured but unrealised assemblages of units. This is illustrated in Figure 3.8. The first of these mappings operates between dissimilar domains (conjectural meaning units in verbal form), the second within one domain (general verbal structure, precise verbal structure).

In design, two similar stages are detectable (Figure 3.9). As in language, each is representable as a mapping, in the one case between dissimilar domains, in the other within a single domain. The first mapping, made possible by the abstract pre-structures, goes from a statement of the problem into a general solution. The second maps from a general solution to a particular solution. The language analogy confirms that the distinction suggested by Alex Gordon between strategic and tactical design is not simply a useful approximation, but a basic attribute of all processes which involves the use of some pre-structured code in an open-ended situation.[13]

A further useful point may be derived by examining this simple basic structure more closely. First, in language as well as design, the mapping from the general

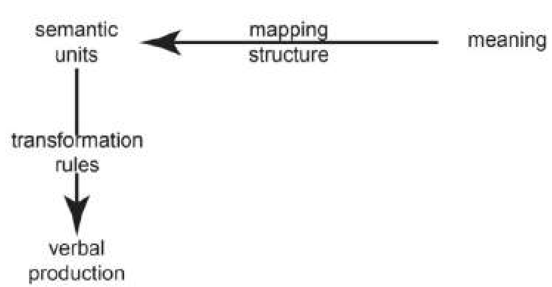


Figure 3.8 The L-shaped theory of language.

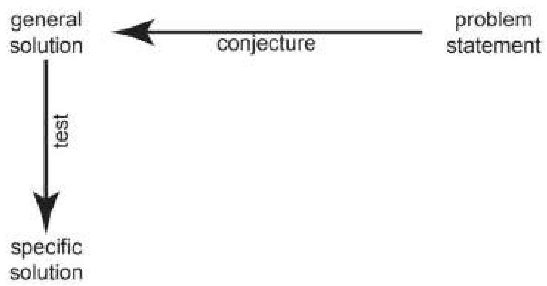


Figure 3.9 The L-shaped theory of design.

solution (semantic units) to the specific solution (speech production) constitutes a test of the realisability or sayability of the conjectured set. Thus, this two-stage process is formally equivalent to the conjecture-test sequence previously hypothesised. Since it is now widely agreed that such a sequence also underlies science, it appears that this simple idea can unify speech, design and science as being transformative of the same basic cognitive activity. In design, the normal way of testing a conjectured solution is to try to specify it more precisely (thus fitting the conjecture-test sequence into the overall programme of variety reduction) and ultimately by building it.

Within this structure, two different strategies exist, although both are variants on the conjecture-test form. The analogy of a detective who has a crime, a set of clues and a set of possible villains may be used. The crime and the villain may be imagined as the end point of a chain, the links of which are the clues in a deductive sequence. The detective may either conjecture a villain and attempt to construct the clues in such a way that they lead to him. Or he may try to build the chain of clues from the crime 'outwards', in the hope that they will lead to a unique villain. Both are conjecture-test, or trial-and-error strategies, and both construct particular actions in terms of the overall logical structure. The same alternative exists for a primitive man who has sticks, flints, slings, spears and twine but who does not yet have a bow and arrow. Whichever way the designer operates, his dependence on the given overall structure and the undiscovered potential of the situation is apparent.

These events in local time are in general of less significance in the overall design of the artificial environment than events that take place, less obviously, in evolutionary time. These changes have the effect of altering the general form of the pre-structures which designers bring to bear on individual problems of phenotype production, effectively acting as a framework of autonomic assumptions within which they operate. At all stages it is the designer's assumptions, the unconscious givens, that provide clues to changes in the morphology of social space. The present time is as good an illustration as any. Our societies are currently evolving new spatial forms which, in ways which are not understood, seem intimately connected to a parallel social evolution. To designers, these changes appear as arbitrary acts of choice. But they will not appear so to anthropologists in the future, any more than the evolution of the structurally stable town form many thousands of years ago appears to us now as the result of arbitrary choices. If design is to be understood beyond the mistaken slight phenotype modification for the continuous recreation of a genotype, then observers must become something like 'archaeologists of the present', attempting to uncover the elementary forms by which our societies are being mapped into new spatial structures. The question 'how does the artifactual environment influence the behaviour of society?' is a manager's question, not the first scientific question. The first question for a theoretical science of the environment is 'what social behaviours determine the environment?'. When research legitimises the second question, and not until then, will environmental science evolve in a non-trivial way. This will not happen while, for example, the application of mathematics to urban structures remains at its current level of aggregating and relating observable collections of arbitrarily selected objects for immediately instrumental aims. Only a model which aims at the generation of socio-spatial form can begin to give the understanding of ourselves that must form the basis of an environmental science.

Such a model is not as inaccessible as might appear initially.[14] For example, many inconsistencies disappear as soon as the local time of design and its object functions are replaced by consideration of the same relationship in its evolutionary time form of a higher-order language constructed out of the socio-logic of such relations. For example, at the object level, the relation of activity to space appears to require some permutation of the concept of 'fit', 'misfit', 'loose fit' and the like, whereas in the higher-order language the relation requires another type of formulation based on concepts of discovery, elaboration and transformation. In fact, the design model outlined above may be continued through the user who, like the designer, takes a pre-structure, elaborates it, discovers it, and cognitively and actively transforms it. This continuous process of transformation, expressing the relation of behaviour to space at all stages, is given in Figure 3.10. Within this model, 'fit' may be assimilated as the efficient elaboration of a particular pattern of activity in a particular pattern of space. But the opposite end of the spectrum – at

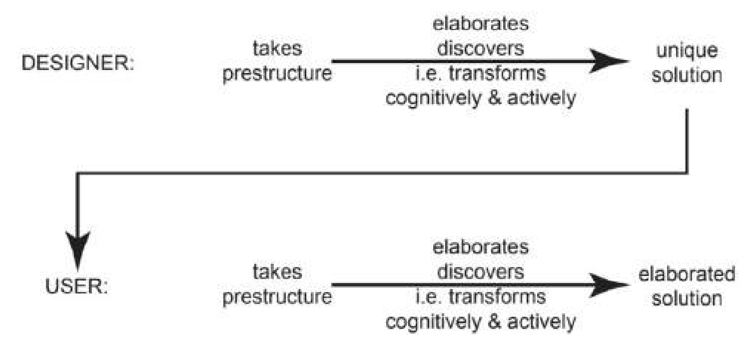


Figure 3.10 The model of the relationship of behaviour to space.

the Camden Roundhouse, for example where a locomotive shed became a theatre – does not appear as an anomaly, or the result of some quite different process. The theory based on discovery, the active form of the relationship, contains both within a single theory.

The principle of elaboration may be used to outline a theoretical approach to space where the fixity of artificial space – the property which confounds the orthodox models – becomes a primary factor. Such a theory begins with the observation that the simplest structures in environmental action are already complex structures.

Such elementary structures, given that they are identifiable, will contain within themselves rules for combination into the higher-order aggregations which give the spatial structure characteristics of urban and other higher-order spaces, as mappings of social processes. These will lead, by further elaboration, to different forms of structural stability to which both social and physical factors contribute. Of particular interest are those structurally stable forms whose stability appears as a function of socio-spatial denseness. Such a theoretical approach to urban systems is in many ways similar to that recently outlined by René Thom[15] as a basis for applying the mathematical theory of catastrophes to biological and cultural systems. Difficult and scarcely scientific as the theory may be at this stage, it is only by such or similar approaches that the broad genotypical changes which construct the individual designer's autonomic framework can be identified, and, in the long run, a genuine liberation from the spatial prisons we are currently building for ourselves be envisaged.

About the authors

Bill Hillier and Adrian Leaman are members of the RIBA Intelligence Unit. The authors wish to thank Hara Drakou, Richard Bennett, Tom Reagan and students at University College London, who have contributed to the formulation and discussion of many of the ideas in this paper.

Notes

1. N. Chomsky, *Syntactic structures* (The Hague: Mouton, 1957).
2. J. L. Elman, 'Finding structure in time', *Cognitive Science* 14 (1990): 179–211.
3. B. Hillier, J. Musgrove and P. O'Sullivan, 'Knowledge and design', in *AIA-Architect-Researchers' Conference. Environmental design: research and practice. Proceedings of the Third Conference of the Environmental Design Research Association* (Los Angeles: University of California Press, 1972).
4. H. Simon, *The Sciences of the Artificial* (Cambridge, MA, and London: MIT Press, 1969).
5. D. Michie, 'Machines and the theory of intelligence', *Nature* 241 (1973): 507–512.
6. F. De Saussure, *Course in general linguistics* (New York: McGraw Hill, 1959) (1966 edition).
7. Editors' note: the word is 'lecton' in the original, but after consultation in the OED and Britannica (https://www.britannica.com/topic/lekton) the editors have opted to spell with k. See The Editors of Encyclopaedia Britannica, 'lekton', *Encyclopedia Britannica*, 5 Dec 2018, https://www.britannica.com/topic/lekton. Accessed 13 October 2023.
8. R. Thom, 'Topological models in biology', in *Towards a theoretical biology*, vol. 3, *Organization and stability*, edited by C. H. Waddington (Edinburgh: Edinburgh University Press, 1970), pp. 89–116; R. Thom, 'Structuralism and biology', in *Towards a theoretic biology*, vol. 4, *Essays*, edited by C. H. Waddington (Edinburgh: Edinburgh University Press, 1972), pp. 68–82.
9. B. Hillier and A. Leaman, 'System, structure and transformation', *Transactions of the Bartlett Society* 9 (1973): 36–77.
10. B. Hillier and A. Leaman, 'A new approach to architectural research', *RIBA Journal – Royal Institute of British Architects* (December 1972): 517–521.
11. T. Winograd, *Understanding natural languages* (Cambridge: Cambridge University Press, 1972).
12. Hillier et al., 'Knowledge and design'.
13. Editors' note: while this mention of Gordon is lacking a citation, Hillier and Leaman were likely interpreting his 1971 Royal Institute of British Architects (RIBA) address, titled 'Architecture: for love or money?', where Gordon introduced the concept of 'Long life, loose fit, low energy,' emphasising the importance of designing buildings that are durable, adaptable and energy efficient. While he didn't explicitly discuss the distinction between strategic and tactical design, Gordon's emphasis on sustainable and adaptable architectural practices aligns with the distinction made here between strategic (long-term, adaptable design) and tactical (short-term, specific solutions) approaches.
14. In the original version of this article, Hillier and Leaman cite it in its previous form, as a paper given at a conference on *Models and Systems in Architecture*, Pembroke College, Cambridge, September 1973: B. Hillier and A. Leaman, 'The architecture of architecture: Foundations of a mathematical theory of artificial space', in *Models and systems in architecture and building*, edited by D. Hawkes (Hornby: Construction Press, 1975), pp. 5–28.
15. While this mention of Thom is lacking a citation, the editors surmise Hillier would have been referring to the following, which existed as a much-thumbed copy in his office well into the 1980s: R. Thom, *Modèles mathématiques de la morphogenèse: recueil de textes sur la théorie des catastrophes et ses applications* (Paris: Union génerale d'éditions, 1974).

4 The architecture of architecture (1975)

Syntactic generators and semantic algebras. Introduction to 'The architecture of architecture'

John Peponis

'The architecture of architecture' is the first published presentation of a theoretical model of built space by Hillier and his colleagues. The central idea is that of a *code*, linking the *syntactic* principles that generate an observable spatial form, with the *semantic* categories that underpin its social meaningfulness.

Consider a simple enclosure and a relation of permeability between interior and exterior. The differentiation of *categories* of space is already implied by the presence of a boundary. The connection across the boundary affords restrictions on access and enables *control*. In addition, rules may specify what can happen in circumjacent space. It is proposed that enclosure can be a gesture of *inclusion* or *exclusion*. The inclusive gesture faces inward and needs not keep surrounding space open. Adjacent enclosures may be built subject to rules of aggregation specifying how many sides an enclosure can share with others, and rules of permeability specifying whether each enclosure should be accessible without going through the interior of another. The exclusive gesture, on the other hand, faces outward and requires that open space be preserved.

Complex patterns arise as simple rules get applied to outcomes already created. Thus, the gesture of inclusion can be realised by aggregating multiple enclosures to jointly surround an open space, such as a courtyard. The fundamental idea that spatial arrangements can be understood as outcomes of a small set of elementary generative principles with an intrinsic social logic was much further developed a year later in the paper 'Space syntax'. It found its clearest expression, with a fully elaborated notational system, in Chapter 3 of Hillier and Hanson's *The social logic of space*, in 1984. But then, semantics, if at all addressed explicitly, are subsumed into the description of syntactic relationships: the social labels assigned to spaces make them non-interchangeable even when they hold syntactically equivalent positions; as a result, a 'longer description' is needed than would be necessary to account for the unlabelled physical form.

With hindsight, therefore, we might ask whether there are avenues of enquiry that might still be of interest if one was to take a fresh look at 'The architecture of architecture'. Note that in classic logic nothing can be simultaneously *a* and *not a*. Hillier and Leaman propose that we can only account for the production of semantic categories through social practice, including the practice of generating spatial configurations, by allowing that the intersection of *a* and *not a* is not necessarily empty. Take the dichotomy between an enclosed space and a space defined as open by virtue of being adjacent to an enclosure; a garden can then be treated as a hybrid which is both enclosed and open. The 'negative counterpart' of such a hybrid might be an enclosed open space set apart and symbolically inaccessible in normal time, for example a sacred space. The idea of a *semantic algebra* proposed in this paper is about accounting for the *types* of space found in society by looking at their semantic construction.

The theoretical rudiments for building such an ambitious model are partly anchored in the distinction between *instrumental* orders, associated with work and production, and benefitting from physically dense patterns; and *symbolic orders*, associated with ritual and requiring the preservation of open space. The former are associated with inclusive gestures while the latter with exclusive. If, however, dichotomies can give rise to new categories, it may be possible to model the range of the different kinds of spaces that characterise a particular culture: streets, squares, parks, sacred spaces, gardens or administrative buildings. Hypotheses regarding such underlying socio-spatial dynamics are perhaps better elaborated in the final chapter of *The social logic of space*. But the ambition to model the semantic categorisation that is realised in spatial syntaxes is no longer in sharp focus.

'The architecture of architecture' is demanding on the reader. As a foundational paper, it introduces and makes explicit the deeper philosophical premises that underpin the proposed models. The structuralist idea that language, cognition, kinship systems and other social phenomena have an underlying formal structure which can be made explicit in appropriate mathematical languages is a key motivator. The proposed models are intended to account for how the social production of distinctive and complex spatial morphologies is possible. They represent the *internal* logic of a morphogenetic system. From this point of view, the epistemological stance taken by Hillier and Leaman challenges and complements the underlying premises of the early performance-modelling papers presented at the same conference.

The guided tour through entropic and morphogenetic time, purely physical systems and systems that contain instructions for their own development, statistical and structural stability, and organic and mechanical solidarity, may appear alternately elucidating or distracting. However, the creative tension between high-level abstract conceptualisation and the articulation of ideas through precise diagrams and symbolic notations gives the paper its intense pulse, the power to breathe into the reader insight already gained and the premonition of further insight yet to be unravelled.

The architecture of architecture

Foundations of a mathematical theory of artificial space

Bill Hillier and Adrian Leaman

Originally published: Hillier, B. and Leaman, A. 1975. 'The architecture of architecture: Foundations of a mathematical theory of artificial space'. In: Hawkes, D. (ed.), *Models and systems in architecture and building*. Hornby: Construction Press, 5–28.

> 'The essential unit entity in biology is not the isolated individual, but is the continuous configuration in space time which connects a parent individual with its descendants (or more generally the union of such configurations relative to species which have among themselves functional interactions . . .'
>
> René Thom[1]

> 'One must force the frozen circumstances to dance by singing to them their own melody'
>
> Karl Marx[2]

Where do intentions come from?

Intentions are a favourite theme in architectural discourse. Not only are they said to be the starting points for design, but also that which distinguishes architecture as an art from architecture as a science. Since science deals with how things are, not how they should be, architectural 'intentions' are said to be the responsibility of the individual designer, or the bodies who instruct him, and architectural science is asked to concern itself with the perfectibility of a process, a methodology for the realisation of pure intentions.

But where do intentions come from? At the very least, intentions must exist by virtue of the prior existence of mental representations of a morphology – that is, some set of related forms – linked to search procedures for the production of design conjectures. These are the structures the designer thinks with rather than things he thinks about, as unconscious as the syntactic and semantic structures of language. The designer's dependence on them is easily forgotten. Too often the concept of 'intention' serves as a means of preserving this amnesia at the level of theory. It locates the designer in a void, rather than in a universe of discourse already highly structured through the history of the morphology on which he acts. It dissocialises the theory of design and renders central questions – such as how the individual acts of the designer transmit essentially social experiences – inaccessible to systematic enquiry.

As soon as the relation between the manifest morphology of architectural forms and the designer's imaginative activity is recognised in theory, the conceptual system of 'intentionality-cum-methodology', with its separation of ends and means, breaks down. It is the essence of any design activity that ends and means are linked into a single conceptual system. Any reference to ends is already resonant with means, and vice versa. It is this unity that makes design possible.[3] The existence of a morphology is the precondition for an imaginative and creative activity. The study of that morphology is therefore the study of the conditions within which design is possible. This is the theme of the paper.

But if architecture-as-activity is only possible because architecture-as-things exists, then there is more than a little inconsistency in arguing first that architecture is unlike science because it is concerned both with ends and means, and second that 'intention' can be separated out as the special preserve of non-science. But this conceals a more important oversight. Science has been concerned with systems which incorporate 'intentional' or 'purposeful' behaviour for nearly two hundred years, more or less since the notion of function became a main theme in the scientific investigation of natural forms. The notion of function as a creator of form, in a suitable environment, led to a new understanding of the importance of time, resulting eventually in the vast new perspective of evolution theory. Architecture, in spite of its predilection for organic analogies, has not yet passed through a parallel emancipation. It continues to operate in three-dimensional space, rather than in four-dimensional space-time. The belief in the autonomy of intention in architecture is by and large an illusion resulting from the removal of time from the morphological field on which architecture acts, and with which it thinks. The re-introduction of time has the effect of making architectural science, potentially, and if we wish it, co-extensive with architecture, and makes it possible to study how the initiatives of the individual designer are linked to the higher-order systems by which each society constructs an artificial environment in its own image.

Physical time and biological time

Two distinct conceptualisations of time, one rooted in physics, the other in biology, have co-existed in science since the last century. The physical view of time developed through the notion of irreversibility. Any reversible process was shown to be accompanied by an irreversible sub-process, which continues in the same direction regardless of the direction of the main process. The irreversible sub-process is the progressive loss of energy available for work. At a general level, this is identified as the tendency to increased disorder, or entropy, in any system. The formulation of the concept of entropy thus led to a physicalist interpretation of time with an arrow pointing in one direction only. This concept might be conveniently referred to as *entropic time*.

Biological time was quite different, and in important ways opposite to the entropy interpretation. Darwinian evolution demonstrated that the variety of observable species was not to be accounted for in terms of the variable effects of the environments, but through the successive transformation of species into new species. Prior to Darwin, 'environmental' determination was the scientific orthodoxy (formerly held by Darwin himself) with transformationism an improbable metaphysics. The transformational time defined by an evolutionary process was not marked by any kind of regular clock, not even the cosmos itself. It required no exogenous 'frame'. It was defined by the successive appearance of new forms, in which the next was possible only because the last had occurred. This can be referred to as *morphogenetic time*.

Both concepts use nature to record time with unique directionality. But there is a difference. Entropic time is a time of observers located outside a system, able to observe changes in relations within the system and call them entropy. Morphogenetic time is time internal to the system, of the nature of its being a system and its form of organisation. It yields a space-time unfolding into greater rather than less organisation, eventually giving, as is the case with biological evolution, systems which are so logically complex that they see themselves, and begin to construct their own evolution.

Through the agency of physical analogue and statistical mechanical models, the physical view has dominated environmental studies. At a theoretical level, this dominance had been reinforced by the adoption of 'general system theory', which attempts to reduce biological (or morphogenetic) time to physical (or entropic) time, in order to gain the advantages of the more easily available calculus associated with the latter. This has led to a generation of non-morphogenetic models which are in effect highly conservative, even leading to the covert introduction of 'social Darwinism', which reduces evolution theory itself to physical time by emphasising the mechanics of natural selection at the expense of morphogenesis.

The loss of the morphogenetic theme is a high scientific (perhaps also social) price to pay for the instrumental advantages of a calculus. But there is another less obvious loss, and that is the possibility of bringing 'intentions' and 'purposes' within the domain of scientific enquiry. The notion of a system which projects a future state of itself by means of some set of 'symbols' (that is, entities that represent other entities, in an informational sense), which do not resemble the projected state as much as act as instructions for building it, is as old as Leibniz, and today offers the general explanatory framework for the study of living organisms at most levels of resolution. It provides a general theory of the means by which nature realises both the continuous production of structurally stable forms, and, in the longer view, morphogenetic transformation. To think of symbolic projections of future states in any system as being solely to do with extra-scientific 'intentional' behaviour (and in architecture to reduce this to the individual imagination) is clearly the result of taking too limited a view of the system, in effect taking it out of its temporal context. What appears at any moment in time as 'intentionality' is itself a possibility, created by the unfolding of that morphogenetic system up to that point in time. To say, for example, that a clock fulfils the purpose or intention 'to tell the time' may be adequate in everyday language, but from a scientific point of view it is misleading. To 'tell the time' is not a construct which exists independently of the artificial morphology defined by 'clocks'. It is a conceptual 'intention' existing only by virtue of the prior existence of that morphology of clocks unfolding in space-time.

The morphogenetic analysis does not of course solve the problem of 'is' and 'should', although it constrains it within a more limited field. But it does introduce a useful new category: the category of the 'might'. Any morphology must be defined as a set of possibilities, not all of which necessarily exist. The actual is not uniquely possible, but an instance of the possible. A morphological set may be thought of as existing either as a distribution of variety regardless of time, or as an ordered, evolutionary series with a unique time direction. It is noteworthy that, at a sufficient distance from our own society, this framework of analysis is normally adopted. An archaeologist or an anthropologist will not see 'intentions' or autonomous 'purposes', but an evolving morphology which links social action and thought to artificial things. The scientific study of contemporary artificial systems – it might be called the archaeology of the present – requires a parallel distancing.

Morphologies and codes

Any science, natural or artificial, has for its object a morphology that exhibits similarities and differences in space-time. The invariances, local and evolutionary, of the set is the object of theory. A theory takes the form of a model, made of symbols, which attempt to represent the space-time morphology as a set, or series of transformations. Operations on the symbols should reproduce, or be mappable into, the observed transformations in the morphology. If this is successful, then the invariances of the morphology might be held to be recaptured in the symbols, and a theory to exist. But to evaluate success, it is necessary to converse with the morphology without the intermediary of symbols, that is, we must displace it physically in some way (experiment), or simply observe it. The theoretical-practical duality of any scientific endeavour follows naturally from these considerations, as does the logical similarity of superficially different types of theoretical activity. For example, in one case the symbolic model may be an equation, the space-time morphology of a certain series of states of falling bodies, and the mode of direct conversation experimentation; and in another the symbolic model may be a series of sentences in natural language, the morphology human societies, and the mode of direct conversation systematic observation. In all cases, however, the results of science are structures of symbols which represent morphologies.

If natural morphologies are already internally structured (that is, generated) through the mediation of symbol-like entities, then it follows that a scientific theory may actually take the form of simulation, in artificial symbols, of some structure of 'symbols' that already appears to exist within nature itself. The 'genetic code' is an example.

If the morphology in question is artificial rather than natural – say, language – then the situation is still more convoluted. First, any invariance in the morphology, which a theory of that artificial system will attempt to recapture, must be assumed to have already arisen through the agency of active and cognitive behaviour on the part of human beings, or their collectivities. Second, it must be assumed that some, or even all, of these invariances are already 'understood' in the same sense that the syntactic and semantic structures of language are understood, and in everyday use, without this 'understanding' being externalisable. Indeed, it is largely through such autonomically 'understood' structures that language permits – to use Chomsky's memorable phrase – a 'rule-governed creativity'.

Many artificial morphologies exhibit these paradoxical properties of being autonomically 'meaningful' and usable while resisting scientific explanation as to how this is possible. Systems which are entirely non-mysterious in everyday life remain more inaccessible to scientific understanding than much of nature. Languages, cities, societies, cultural systems, architecture, even knowledge itself all illustrate this paradox. As a consequence, special sciences have come into existence to study these systems, namely the sciences of the artificial. The branch known as 'structuralism' explicitly studies what is already 'known' in the autonomic sense; that is, it tries to reconstruct the conditions in which such 'knowing' is possible.

The morphologies that exhibit these paradoxical properties have one attribute in common: each either is, or contains, a system of social signification. It is this, rather than the physical nature and form of the morphology, that renders it opaque to present scientific understanding. In the case of architecture, the problem resides in how architecture 'works' in relation to society. How and why do different societies at different times construct and transform artificial space, and what is the effect of these transformations? How can the *intelligibility* of artificial space be accounted for and reconstructed?

The concept of 'code' is useful in analysing such morphologies. The word 'code' has two slightly different meanings. It can be a structure of rules governing a particular morphology, say, a set of permissible behaviours. Or it can be a structure of rules for making translations, either between differently structured domains of the same morphology (as in the case of a code for transmitting secret information) or between two dissimilar morphologies (as in the case of a morphology in space-time related to an abstract or semantic morphology – language for example). The latter is the important one from the point of view of artificial morphologies which serve as signifying systems, since the formal requirement for inter-subjective intelligibility to be available through such a system is that transformation within the space-time morphology should be systematically related to transformations in the semantic morphology. This would be the minimum structure necessary for such a system to transmit meaning.

If, in addition to the intelligibility requirement, the further stipulation is added that the morphology should work as a 'rule-governed creativity' – that is algorithmically, rather than on a basis of one-to-one mapping – then it implies that each of the dissimilar morphologies related in the code must possess its own internal structure, and that the translatability between the two must be at a level of parallel structural transformations, and not at the level of looking up forms in a dictionary which assigns a unique 'meaning' to each. Language exhibits this property in both its domains. Words tend to be organised into syntagms and syntaxes. The 'semantics' embody, at the least, a bundle of interpretive schemes for dealing with reality and capable of 'solving problems' introduced by unfamiliar experiences. Structures of this type are the subject matter of categorical algebra, and the best way of representing them formally is by treating the morphologies as the domains of an algebraic structure. A simple, yet mathematical, interpretation of the concept of 'meaning' is yielded through such formal structures of relationships.

Lévi-Strauss has written extensively on the algebraic bases of cultural signifying systems and has ascribed their existence to inherent properties of the human mind. The foregoing considerations suggest that his analysis is unnecessarily extreme. If account is taken of the 'dissimilar domains' structure of such codes, and the 'internal structure' requirement, then it is clear that during the evolution of the system there would be a need at each stage to maintain translatability between domains, and this could only be based on a formal translatability between the – possibly different – structures within either domain. In other words, a need for an algebraic mappability would ensure that at each stage of evolution, the whole system, and each domain separately, would retain an algebraic form by being related to the other. The space-time morphology would be internally ordered through the evolution of structurally stable states as it moved from simpler to more complex forms. The internal code of this morphology would be an abstract deposit of these structurally stable states. The concepts of structural stability and morphogenesis are keys to the nature of artificial codes, and translatability between dissimilar domains can be suggested as an alternative general theory for the algebraic nature of codes.

The concept of code thus offers a framework for the scientific analysis of artificial morphologies that is both rich and rigorous. It links together, in a natural way, aspects of the subject matter that are normally well separated and even thought incompatible. For example, it links the social and historical aspects to the mathematical. It links the study of modes of creating artificial space to the study of its intelligibility. It links the notion of underlying formal structure to that of a rich variety at the observable level. It links existent and past states to possible and non-existent states. The code is the means by which the unfolding of a morphology in space-time 'passes through the head'. By studying the history of that unfolding, and in particular its structurally stable states, in relation to the societies that produced them, we are in all likelihood studying the means of our own awareness of space, that is the cultural intelligibility of the morphology of artificial space.

The analysis of an artificial system on the basis of the concept of code requires three related studies. First, the study of a morphology and its internal rules of evolution, giving what might be termed an absolute or 'mindless' unfolding of that morphology. This, for convenience, may be called the syntactic domain. Second, the study of a semantic field to which the morphology refers, in a signifier-signified relation. This may be called the semantic domain. And third, the study of the translatability relations. The concept of code might be held to refer simply to the last, the translatability relations, but since these are meaningless without the two domains, then it is better to refer to the whole structure only as the code.

Statistical stability and structural stability

Science, being concerned with the introduction of novel artificial structures and relationships, namely those between symbolic systems and morphological domains, takes the form of a dialogue between the formal capabilities of symbolic systems and the morphological domains they aim to encapsulate. This requires the scientist to be aware of the *constructive* effect of symbolic systems on the world they describe. The universe seen through symbolic systems is likely to reflect the internal properties of that symbolic system. This can happen in mathematics as well as natural languages. Social scientists, for example, have mistaken the forms of the calculus, with its central notion of continuous concomitant variation, for mathematics itself, and even for science itself. Only relations, however few, reflecting this structuring effect of the dominant symbolic system were permitted to be 'scientific'. The use of structural mathematics in related fields like anthropology, and the growing knowledge of catastrophe theory,[4] are only slowly leading to an awakening from this 'dogmatic slumber' of the social sciences.

Awareness of this danger suggests that it is useful to establish a metacritique of the types of formal language available, and the properties of morphologies to which they might refer. In general, three types of property that morphologies present can be distinguished and referred to the properties of symbolic systems. First there are *logical* properties (for example: 'existing', or 'not being something else') which are, quite simply, the minimum properties of all objects. To this domain corresponds the symbolic system of logic, which deals with those properties of objects which they share with all other objects. Second, there are *measurable* properties, the separating out of which was Galileo's great contribution to theoretical physics, that is, his 'primary' properties, as opposed to the 'secondary' properties which were imposed upon the object through our mode of perception. To these correspond the mathematics of quantity. Third, *semantic* properties, namely those which result from significant connectedness within and among morphologies, a subset of which map into human cognition, and are called 'meaningful'. To these correspond the mathematical structures we normally think of as algebraic. The relations among, and mutual reducibility of, these different types of mathematical structure are of course of great importance in the history of epistemology, but they are not our theme here. Here this distinction is introduced in order to clarify the role of statistics in modelling artificial systems.

Statistics is essentially the logic of aggregates. Like logic, it requires only that a set of objects has a property, say, of corresponding to a point on a continuous scale, or of adopting, at any point in time, one of two fixed states. Then, given that the process is random, meaning that there is no connection between objects or events (no significant connectedness, or semantics), statistics can make good predictions about the behaviour of such aggregates, since in large enough samples these will approach statistical stability as expressed, for example, in a normal distribution curve. Where the process is non-random, for example when the events in a sequence are such that an event is influenced by previous events in the sequence (as in a sentence for example), then we introduce more algebra into the equations.

The importance of this randomness requirement is not always realised. It implies that when a set has a strong internal structure or is observed only in small numbers, then its properties or behaviour will less easily conform to the basic pattern of statistical stability. Structure is the enemy of statistical stability. To use a simple example, if we regularly patronise a coffee house, and the strength of the coffee varies from being very strong to very weak and back again, rather than oscillating around a mean strength, and this holds even for a large number of visits, then we know that the process is non-random, and that some stable structure has interfered in the system to produce statistical disorder – that is, a stable deviation from the normal distribution curve that a random process of coffee-making would produce. For example, there exists a rota for coffee-making on alternate days and the different coffee-makers make it wrongly in a different way. This relation between structure and statistical disorder at the level of resolution at which we first seek it is of the utmost importance. Moreover, for large numbers, it is essential that the interference structure should be structurally stable, or it would not produce interference in the statistical stability of the sample.

This quickly leads to qualitative break-off points in the type of order we can hope to find in a system. Consider a simple spatial example of swimmers depositing towels on an undifferentiated beach. The order in such a system could be described, for large numbers, in terms of some probable spatial distribution of clustering, taking into account such factors as point of entry to the beach and so on. But take those swimmers away from the beach and put them in a more complex environment, say, wooded dunes, and give them a packet of sandwiches and a qualitatively different situation appears. In this case, the distribution will be largely determined by a new stable structure, namely that the swimmers will locate themselves by finding realisations in the rich environment for a model 'in the head' which tells them what a picnic place is like. Although we have only increased the degree of differentiation in the system, it is clear that our concept of order must be of a different kind from the statistical one that worked in the less differentiated case. The situation is still highly ordered, but not in terms of a statistically stable distribution on a surface.

In social science it is these structurally stable deviations from statistical stability that should be studied. Statistical disorder is interesting because it requires us to look for such a structure. In this case, the possibility of a purely statistical order acquires a very useful role as a starting place, since it tells us what the system would be like if it were random, that is if it lacked semantic properties, that acted as structurally stable interference effects in the statistical stability. The task of any science is to arrive at a random process in respect of its morphology, that is, to name and relate all structure that produces non-randomness, such that what remains in the space-time manifestation of the morphology is random and therefore epiphenomenal. Codes are structurally stable producers of non-randomness.

Function and form in architecture

That buildings are 'functional' and 'meaningful' is self-evident, but it is equally self-evident, through the morphogenetic analysis, that 'function' and 'meaning' exist and are intelligible by virtue of the evolution of a morphology of built forms. This requires us to make a further fundamental distinction in the study of artificial systems, one which is familiar in linguistics but not in sociology or architecture. This is the distinction between the morphology itself, as it exists at any point in time, and the individual or corporate appropriation and use of that morphology. The first corresponds to what the linguists call the level of language, which is a structure of social conventions independent of the individual; the second corresponds to speech, or the individual appropriation of the existing morphology. The natural linkage of the higher-order, supra-individual morphological domain with the individualised but still structured appropriation of it that linguistics proposes works well for architecture, and leads away from the mystique of 'intentions' without sacrificing creativity.

A conceptual difficulty arises with architecture because the morphology of forms exhibits what we might call 'negative redundancy' in relation to function: that is, nearly everything functions in more than one sense. This can be understood by looking at the evolution of built form in terms of a fundamental duality: that buildings at once function as modifiers of the relation between man and nature and as modifiers of the relation between man and other men. But this duality is not strange. In fact, it is normal. All systems by which mankind changes its relation to nature are also elaborated into systems of social signification. We might consider eating as an example. Eating fulfils an energy transfer need, but at the same time it is embedded into complex and pleasurable cultural forms. In social terms, the functional nature of all such artificial systems is to exhibit this duality. To restrict its function to the physiological aspects desocialises the morphology and renders it unintelligible.

The matrix of these dualities, man-nature and man-man, and 'language' and 'speech', constructs the four-function model of architecture.[5] At the level of 'speech' or 'design', the man-nature relation is the climate modification function of the building; at the level of 'language' or 'morphology' it becomes the resource modification function. The man-man relation at the lower level is the level of interaction between activity and space; at the higher level it is space as a social language constructed through the properties of the morphology. This paper is concerned with the last named, the man-man relation at the level of the 'language as a whole'.

In all such artificial systems the existence of the language-like morphological level is the precondition for the speech-like 'design' level. To try to consider design independently of its morphological domain which makes it possible is like trying to explain speech without understanding the structure and properties of language. The important question is: *how* is it that higher-order morphologies are constructive of speech or design acts? Is it simply by providing a set of morphological images which can be transmitted? Or is it important that, in architecture as in language, this relation should work algorithmically, and not on a one-to-one basis, in which case we must assume it has a more abstract nature? Again, the concept of code comes to our assistance. One aim of the second part of this paper is to show how and why the transmission is algorithmic and abstract, rather than imagistic and one-to-one. In explaining this we shall use the notion of code to link sociological and logico-mathematical properties, and in effect to give a sociological and historical interpretation to the formal properties of built form which are embedded not so much in the morphologies as in the relations defined by morphologies. It is the abstract nature of this transmission structure – the code – that links the generation of built form to our understanding of it. It passes through the mind both ways.

The hypothesis is that the relation of human behaviour and cognition to artificial space is governed by an evolutionary code, which explains both our modes of constructing space and the way it is intelligible to us. The morphology therefore exhibits the properties of social language. The abstract code will account for the similarities and differences in the morphology, and show how our understanding of them depends on a matrix of socio-spatial abstractions. From an empirical point of view, the code can be seen as a structurally stable interference in an otherwise random statistical order, existing at the level of interaction between activity and space. In terms of architecture as a whole, the analysis is incomplete, since it does not deal with the man-nature side of building. Nevertheless, it offers a framework for integrative analysis of both aspects.

The logic of space

The research programme implied by this theoretical perspective is therefore concerned with the reconstruction of a code which produces structurally stable differentiations in the forms of artificial space, and can be thought of as equivalent to a theory of that morphology. The test of such a theory is its ability to map observable transformations into transformations within the symbolism in which the theory is embedded.

This is an orthodox, hypothetico-deductive scientific strategy, but its epistemological stance is quite different from that of system theorists. The meta-theoretical statements of system approaches normally argue that all models are relative to the observer's intentions and purposes.[6] This is not self-evidently true. A covert anthropomorphism underlies the relativistic modesty of system modellers. Surely nature itself has its own models? The genetic code is not relative to observers, since epistemologically speaking we cannot doubt the existence of some such structure. The genetic code may be thought of by us as relative to our sign systems, and the conceptual schemes they imply. But from a scientific point of view such structures exist in nature as 'internal' models, and much of modern biological research is concerned to reconstruct them.

In the domain of the artificial, these concerns are reinforced. Language may be considered as an example of a structure which undoubtedly has a formal structure independent of observers (indeed it is that which gives it its creative usability by individuals). A model for a grammar may be wrong, but this does not imply an epistemological relativity. The difference between 'structuralist' and 'system theoretic' approaches in the sciences of the artificial is given by this distinction between what might be called 'internal' (structuralist) models and 'external' (system theoretic) models; this in turn links into the distinction between morphogenetic time (internal models) and entropic time (external models).[7] The research strategy implied by the internalist programme could be characterised approximately as treating reality as the printout from a computer, and trying to deduce or reconstruct the programme.

Another way of looking at it is to suggest that the Galilean hypothesis (that physical reality *is* mathematical, and that equations are more than useful instruments) is being transferred from the natural to the artificial world. If artificial reality is indeed logico-mathematical in some sense, then it should not be thought surprising. Mathematics and logic are themselves artificial, and since they attempt to deal with the formal and abstract properties that are not written on the surface of things, at the same time as representing our cognitive operations on things, then it would be no surprise if foundational ideas in mathematics were also pervasive structuring forms in artificial domains. Both in biology and artificial sciences, structuralism is founded on this anticipated isomorphism.

In the case of artificial space, the point acquires force through a simple analogy between spatial and logical forms. The simplest gesture of artificial space is that of defining a barrier. A barrier distinguishes two types of space from each other. If no distinction were required, then no barrier would be required. This corresponds exactly to the basic gesture in the logic of classes, that is drawing a distinction by putting a barrier between concepts. Moreover, in space as in logic, this gesture can be made either in terms of distinguishing two types of space both with positively different properties, or in distinguishing one positive type and its complement or negation, that is lacking the property. Again, by making a barrier permeable, we define a relation between types of space. The clarity introduced into formal logic by Venn diagrams might be thought of as illustrating this close analogy between space and logic. Finally, it could be argued that, like logic, architecture begins with simple spatio-logical gestures in three-dimensional space, then elaborates them into many-dimensional structures in logical space. It is through this that in building walls and holes we build intelligibility and complexity.

Spatial surfaces and aggregation modes

In all artificial systems which construct inter-subjective 'meaning', the question arises as to the level of resolution at which this meaning is principally decipherable. In language, for example, it is *not* the level of the word. Any word is prolific with possible meanings, as any check with a dictionary will show. Language nevertheless manages to achieve unambiguous definition of the meaning of a particular word by extensive use of the context of the word, including the other words around it, the epistemological context of the speech act, and such extra-verbal cues as style and gesture. These are not gratuitous and decorative aspects of language, but technical devices for fixing meaning across a semantic field that is vastly larger than it would be if words and meanings were in one-to-one correspondence.

From the point of view of words, the sentence is the minimum viable meaning form, although even this still requires all the context variables. This appears to have an interesting correspondence with the way in which people and architects experience artificial space. Architects both design buildings and tend to see buildings as the basic unit of environmental experience. The architectural culture can be virtually defined in terms of the tendency to select a particular building from an aggregate and give it unique attention. For the layman, unless he shares the architectural culture, the situation is usually different. His experience is usually at the level of 'place', that is the general properties of the location and situation in which he finds himself.

Given that buildings, like words, may be rich with possible meaning in themselves, another promising analogy with language suggests itself. Everyday experience of space appears to be much more at a level comparable with the sentence, that is the aggregate of buildings in a particular location. This is confirmed by the way in which spatial terms and their associated social semantics are embedded in the unconscious semantic structures of language, in which words like 'village', 'suburban', 'urban' and so on are spatial descriptions which carry with them a mode of social existence. They appear to constitute, as it were, islands of structural stability in the socio-spatial code. No term exists in architectural discourse for this level of experience, and certainly no technical term. Since it appears that the elementary level of environmental experience may be that of the 'sentence', or characteristic aggregate, a new term is needed. The term 'spatial surface' is suggested. The primary level at which the socio-spatial code expresses itself through the language of the morphology of artificial space is this 'spatial surface'.

It might appear that in moving up the scale of complexity, an insuperable problem of formal description is likely to arise. Are spatial surfaces not infinitely variable? In certain ways they are, of course, but in terms of certain structural properties they are much simpler to describe than buildings. As is normally the case with hierarchically organised systems, micro-state complexity yields a macro-state simplicity. In some fields it is very difficult to say how this happens (picking up a cup of coffee, for example, is a simple macro-state gesture involving phenomenal micro-state complexity), but in the case of artificial spatial surfaces it is relatively simple. Their general structural properties are given by the way in which the spatial surface is generated. These processes of generation can be termed: *aggregation modes*. That aggregation modes take relatively few basic forms can be seen by considering a little more closely the logic of spatial elaboration.

Buildings are normally thought of as enclosing space. But they also define external space. They do so individually, but more importantly as aggregates they define higher-order spaces. An aggregation mode is a rule structure concerned with what can and cannot be done after space has been enclosed, both in the substantive proximate space, and in the marginal spaces defined by the shape of the building. It is these socially defined rule structures which give rise to the characteristic structural differences in spatial surface, and it is at this level that the morphology of artificial space acts as a social language. Aggregation modes, it will be seen, are invariant under changes of style, technology and decoration. They provide the rudimentary schemes on which decoration may elaborate. It is aggregation modes which account for the manifest homology of socio-spatial forms distributed through different cultures in time and space. Also, the major transformations in the morphology of space that are occurring today are through the dramatic effects of changes in the aggregation modes, and their social semantics.

It will be noted that a radical change in the mathematical strategy is implied by these concepts. A generative approach is to be substituted for a descriptive one. This is a major simplification, since it is argued that *the structural properties of spatial surfaces on which 'meaning' (and therefore the usage code) hangs are given by the generative rules for that spatial surface*. This generative approach applies both to the local (in space and time) generation of a spatial surface, and to the evolution of the code itself on an evolutionary scale.

The concept of system also changes with this approach. The task is not to describe a system as it is presented to an observer, in order to introduce better external controls on its further development, but to understand *how the system generates itself, and produces its own internal order*, that is, its structurally stable states and processes within an evolution of a socio-spatial morphology. In such a system there are no simple elements, only relations between morphological and semantic domains. Moreover, it has already been shown how complex a barrier is, seen from a logical point of view. The secret of the evolution of such systems, and that which solves the problem of relations between micro-states and macro-states, is that the 'simplest structures' (rather than 'elements') of such a system contain within themselves structural variables which lead, when simplest structures are combined together, to variable unfoldings into higher-order systems. It is the object of the next stages of the paper to argue that these unfoldings possess stages of structural stability in both morphological and semantic domains of the code, and in the translatability relations.

The 'system' described is, of course, a large, incomplete and evolving abstraction, which is not directly observable (any more than the equation for a falling body is observable); moreover, it is probable that no complete embodiment of the 'system' exists on the ground, but that nevertheless it is the means by which what does exist on the ground, with all its local and historical contingency, is intelligible. At this stage it is clear that a virtual reversal of the concept of 'model' is involved, since it is argued that this abstraction is the most 'real' thing, and that the realisation in space-time – the artificial spaces we see – are 'models' for this abstract structure of relations. But such a reversal is familiar. This is the concept of model used in mathematical logic, where a 'model' for a structure of logical symbols is a realisation in some mathematical domain.

This might be thought an intellectual extravagance, were it not that this coincides precisely with our present understanding of how language works, and how it relates to acts of speech and understanding. It appears that we need a paradigm in which the whole of language and its semantic field is in action with every linguistic act. It has, for example, been demonstrated that in order to decode certain sentences exactly – which in practice are not difficult but which a computer cannot manage – it is necessary for the decoder to have access to virtually the entire field of knowledge of the speaker.[8] If artificial space is intelligible, there is no reason why it should not be because a knowledge of the whole code is embedded in our representation of the morphological field, such that as with language, the whole structure can be set in action in interpreting environmental experience. How this relation of the whole code structure to the particular situation and experience is constructed will also be part of the subject matter of the ensuing section.

Elementary syntax of spatial surfaces

The general form of a code is that two dissimilar, internally structured domains are linked by translation rules, such that transformations within one domain can be mapped systematically into transformations in the other. Where one domain is morphological and the other abstract, a syntax-semantics distinction can conventionally be used. This duality of a morphological reality and a domain of abstract 'meaning' characterises artificial space, but it is also necessary to make a further subdivision within each domain into two further subdomains, giving four domains in all.

It has already been observed that the simplest architectural gesture, the indication of a barrier, yields an already complex structure, namely a relation between an enclosed and non-enclosed space formed by a barrier, both being mapped into a semantic domain. The four domains of the code are all present in this structure:

1. logical differentiation of types of space, that is the reasons for enclosure or non-enclosure
2. the barrier arising from (1), the act of differentiation, forming enclosures and defining non-enclosures
3. the permeability of enclosures in relation to non-enclosures and other enclosures
4. rules governing what may happen next in the proximate spaces defined by the enclosure.

The hypothesis is that each domain has an internal structure, such that its transformations can be mapped into transformations in all other domains.

Call logical differentiation *A*, barriers forming enclosures *E*, permeability *P* and rules for further elaboration *F*. For purposes of initial exposition, let each of *A*, *E*, *P* and *F* adopt one of two states: the positive state, *A*, *E*, *P* and *F* and the negative state *A′*, *E′*, *P′* and *F′*. Thus, *A* or *A′* distinguishes a reason for being *E* or *E′*, enclosed or non-enclosed, implying *P* or *P′*, the form of permeability between enclosed and non-enclosed (or enclosed and other enclosed) and a rule *F* or *F′*, which says whether or not a further enclosure can be constructed in the proximate space to the enclosure. This gives an overall structure: $[[A \vee A'] \rightarrow [E \vee E']] \rightarrow [[P \vee P'] \rightarrow [F \vee F']]$. This represents the logical form of the simplest architectural gesture. Conventionally, *E* and *P* may be thought of as syntactic, since they refer to the ordering of physical entities, *A* and *F* as semantic since they refer to ordering of abstractions.

For the sake of simplicity, assume that a complete enclosure is made of four barriers in the form of a square. This is the basic unit of operation and carries with it all the structure in the above expression. Let *A* be such that for *each* barrier a rule *F* or *F′* governs what may happen in the external space beyond the barrier, namely *F*, 'free to enclose again', or *F′* 'not free to enclose again'. Thus, for each *completed* enclosure, *A* determines a function *f*, which may have a value from 0–4 according to how many of its barriers may be built against. Thus, for an enclosure with f_0 none of its four barriers may be built against; an f_1 enclosure may have one of its four barriers built against; and so on up to f_4, which may have all its barriers built against. A space, defined by an enclosure, is assigned, through the act of enclosure, a rule which says whether or not the proximate space is free to be built upon or conserved. A randomised process is then assumed to operate, using these rules and structures, in a finite, isotropic space until the space is as densely filled as possible subject to the rules and structure.

Examples of the outcomes of such operations are given in Figure 4.1. Obviously for f_0 the result is a series of free-standing pavilions because each time an enclosure is completed no wall may be built against any other. The surface is full when no more enclosures can be fitted in without touching each other. For f_1 the results are rather similar, but the units are semi-detached; f_2 is more interesting – the more familiar constructions of streets and squares as well as a variety of zig-zag shapes are derived; f_3 is more maze-like, constructing a series of courtyards, usually some of unit size, and some of twice unit size; f_4 eventually fills the whole spatial surface. Each of these is topologically distinct, f_0 is a ball; f_1 two joined balls; f_2 is a line including a line joined to itself, that is, a circle; f_3 is a kind of net; and f_4 is a dense block. It is interesting to compare these results with those of March on the floor-space index of different configurations.[9]

All this is straightforward but illustrates a point that may not be so obvious. A relatively random process of aggregation, given that there is a simple rule, can produce complex, recognisably different and structurally stable higher-order aggregations.

Matters may be complicated by introducing a further factor, namely a permeability constraint *P*, which says that all defined spaces shall be accessible to each other and from the outside. This has the effect of making the spatial surface appear more lifelike. When *P* is applied to f_3, for example, the spatial surface generated is very similar to such structures as the squatter settlements of Delhi and the Chalcolithic village at Teleilat Ghassul (Figure 4.2). The typical sub-surfaces (e.g. two short parallel lines, a U-shape, an isolated block) of these real forms are generated by the formulas $f_3(P)$ operating in an otherwise random process. These simple examples illustrate the idea of *a model in which the mode of generation is said to be the model's structure*.

Such structures can be termed the first-order unfolding of the syntax. They offer an initial simplified approximation of some of the properties that we might wish to characterise as those of spatial surfaces, and show these differences to be, from a simple-minded and purely mathematical point of view, a set of transformations. This may be more theoretically interesting than it is lifelike, because only part of the total code has been utilised, and also because it has only

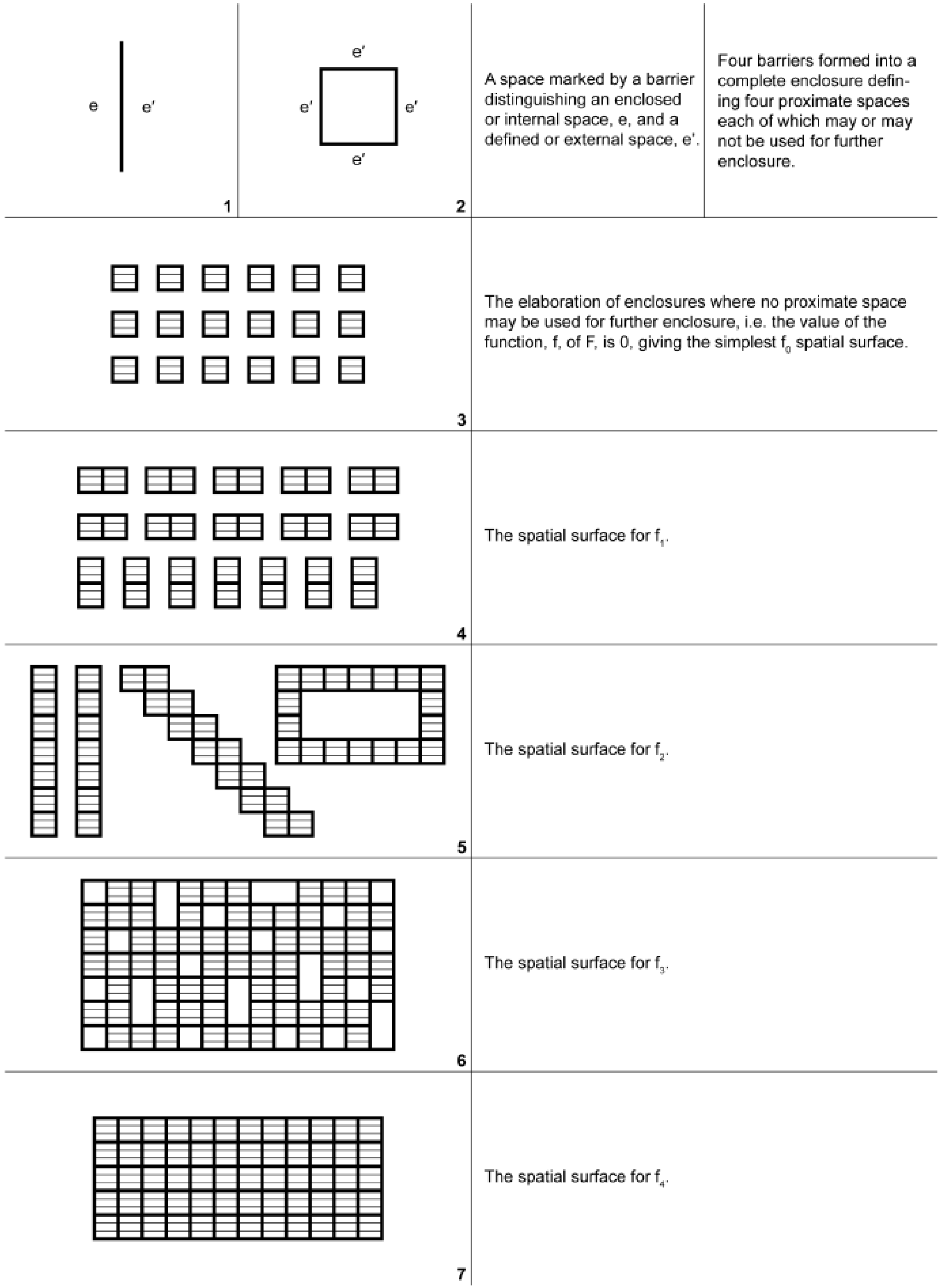


Figure 4.1 Unfoldings of *f*.

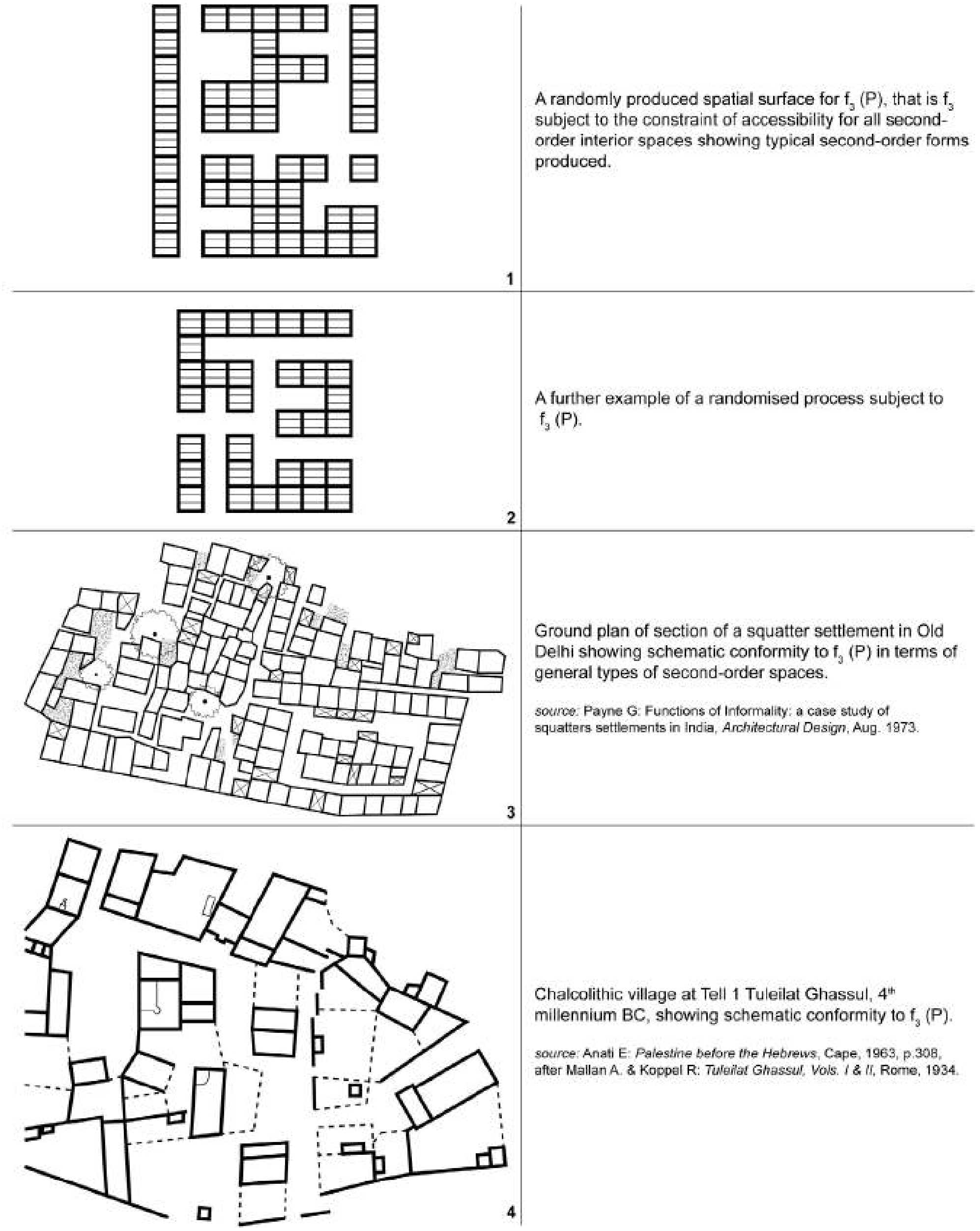


Figure 4.2 Unfoldings of f_3 *(P)*: formal and real.

been applied at one level. The spatial surfaces generated would go on endlessly unless stopped by an arbitrary barrier, or some rule for making finite, structurally stable higher-order structures out of the unfolding of the syntax. The next stage of the argument therefore concerns the production of structurally stable higher-order unfoldings, which we will call the 'elementary forms' of artificial space, since this is what we believe them to be.

The rule for the generation of elementary forms out of the syntactic process for the production of spatial surface is close to the heart of the argument: it is that once a rule has produced a structure, the rule can then be applied to the structure itself. In other words, the rule operates continuously, and it can make itself more complex objects to work on. This is the key to the syntactic evolution of spatial form. In fact, this rule has already been covertly used in moving from barriers to enclosure, that is, using barriers to make the simplest closed form. When four barriers make an enclosure, rules can then be applied to the enclosure as a whole. This operation of 'same rule, more structure' will also be useful in the discussion of the semantic evolution of artificial space. It is the key to the generation of extremely complex, but structurally stable, forms out of simple operations, and shows in principle how the code both produces and renders intelligible the highly complex forms of space that we experience in everyday life.

The next problem is therefore to define what new structure has been created, and what is the effect of continuing to apply the rules in a process which includes the new structure as well as the old. Most of the new structures are self-evident; the set comprises variously connected aggregations of enclosures, some discontinuous, some continuous in varying degrees. These are shown in Figures 4.1 and 4.2. One more new structure is less obvious. Whenever there is a defined space which is non-free, that is, one that is conserved empty, there exists what could be termed an 'open space barrier', which is of indeterminate depth. A secondary, or notional, barrier has been created at some distance from the barrier which defined the space, which exists whether or not it is fixed or marked on the map. In the case of an f_0 structure, this open space barrier extends all the way round the enclosure, and by implication, a more complex structure has already been created than appeared to be the case.

If this notional barrier around the f_0 is completed by being fixed, or even built as a physical barrier, then one of the two major elementary forms had been created, namely the pavilion surrounded at a distance by a secondary barrier (Figure 4.3$_1$). This is termed the 'inout' elementary form, since the enclosure governs a set of open spaces in a relation that is both in-to-out, and one-to-many. That this form has an inverse is well known to all who have seen the motif of Land Use and Built Form Studies.[10] The set of continuous enclosures around a perimeter, enclosing an open space in an out-to-in, many-to-one relation can be termed the 'outin' elementary form. Its generation in the syntax is equally simple. It is the unique product of the rule: maximising the value of f at the first-order level, use whatever structure exists to make a second-order enclosure (Figure 4.3). By contrast, the rule for the inout

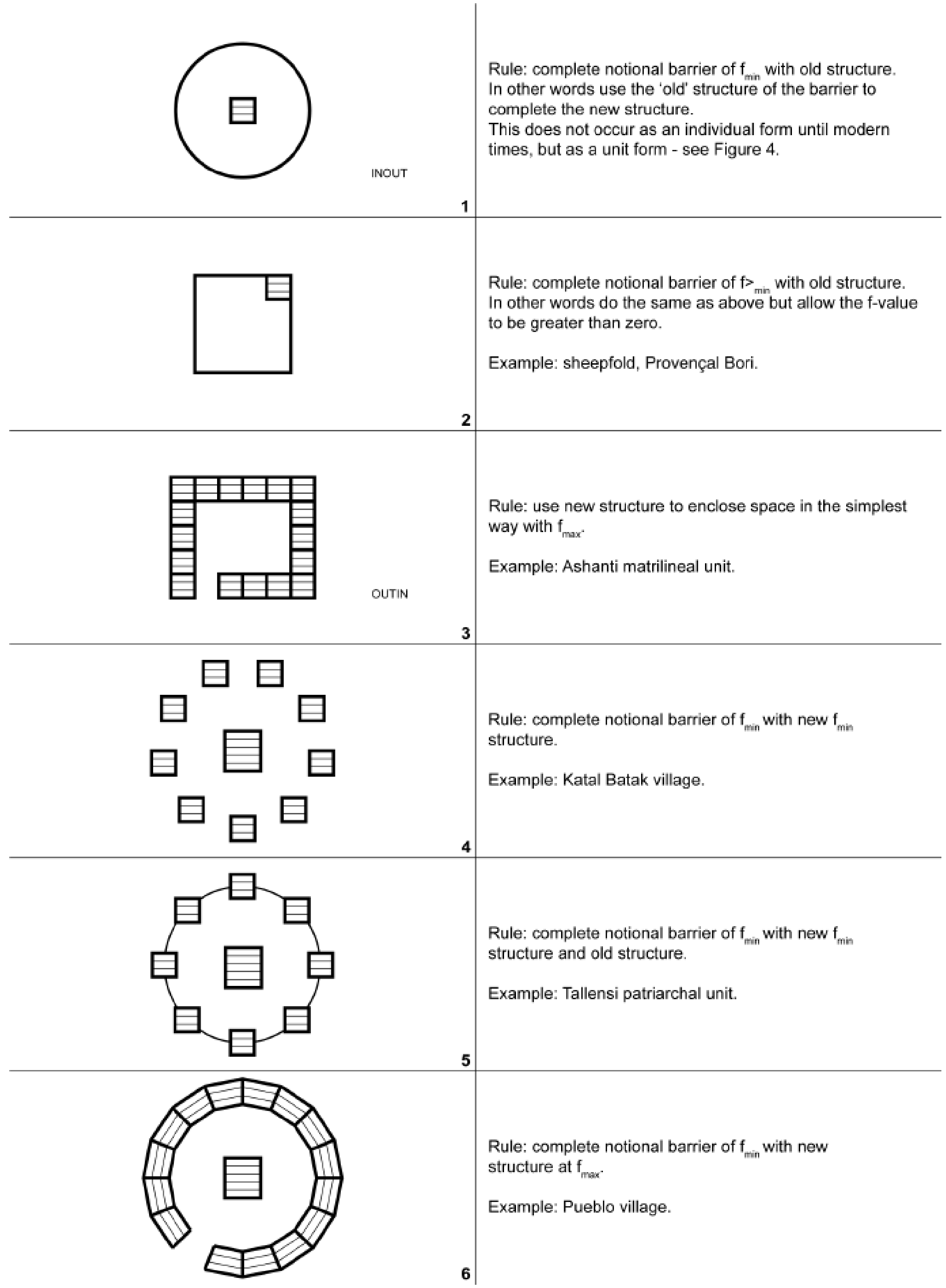


Figure 4.3 Examples of elementary forms.

form is: minimising the value of *f* at the first-order level, use whatever structure exists to make a second-order enclosure. Thus, the syntax produces these two primary elementary forms literally as inverses of each other. It is noteworthy that both forms have a 'double boundary structure', that is one in which two boundaries must be crossed to move from 'outside' to 'inside'. The generation of double boundary structures is of pervasive importance in the evolution of the urban spatial surface.

They are, however, only the two simplest outcomes of the second-order generative process. Figure 4.3 illustrates a further selection of elementary forms that may be generated at the second-order level, together with their generative rules. So far as can be seen at this stage – that is prior to any exhaustive programme of generation and testing against real-world data – these unfoldings offer a syntactic taxonomy for *P* large proportion of village and semi-settled forms that have become familiar in recent years through the ethnographic and photographic works of Rudovsky, Fraser and others.[11] A full interpretation of these villages as socio-spatial systems, however, requires also the semantic component of the argument (discussed later) as well as a study of local and historical contingencies of the normal technological and environmental kind.

A further distinction is critical to the further evolution of complex spatial systems. This is the distinction between what can be termed multiple and unit forms. A multiple form is one that is generated as an aggregate following a rule. Once it exists, however, it is clear that it can itself be a template for further similar forms, without the process of generation needing to be repeated every time. Again, this idea has been implicitly used between the level of the barrier and that of the enclosure. The unit form is therefore a multiple form used as a template. An example would be the inout elementary form used as a template for a large institutional building. A section of more complex unit forms is given in Figure 4.4. All these forms exist in Banister Fletcher,[12] except one which exists only in multiple form, and which is crucial to the evolution of urban form.

The distinction between unit and multiple forms is of fundamental importance in the evolution of higher-order structures, since it permits a design-like activity to be an intrinsic part of the evolutionary process. In fact, it is tempting to view design – when seen within the scope of an evolutionary morphology – to be precisely to do with the use of templates, as opposed to the auto-generation of form following rules. A designed form is of its nature defined as finite, as opposed to self-ending as is an auto-generative, rule-governed form. The consequences of this for the design of a town are interesting, since it appears to suggest a third category between designed order and contingent chaos: namely that of an auto-generative order of a process subject to generative rules. The form of towns as we know them is certainly of this kind. A paper is in preparation demonstrating the syntactic evolution of urban form, showing that it appears, in its familiar forms, as a fourth-order unfolding, a conclusion which is paralleled in the semantic unfolding. This offers a natural and internal theory of the equifinal production of urban form by

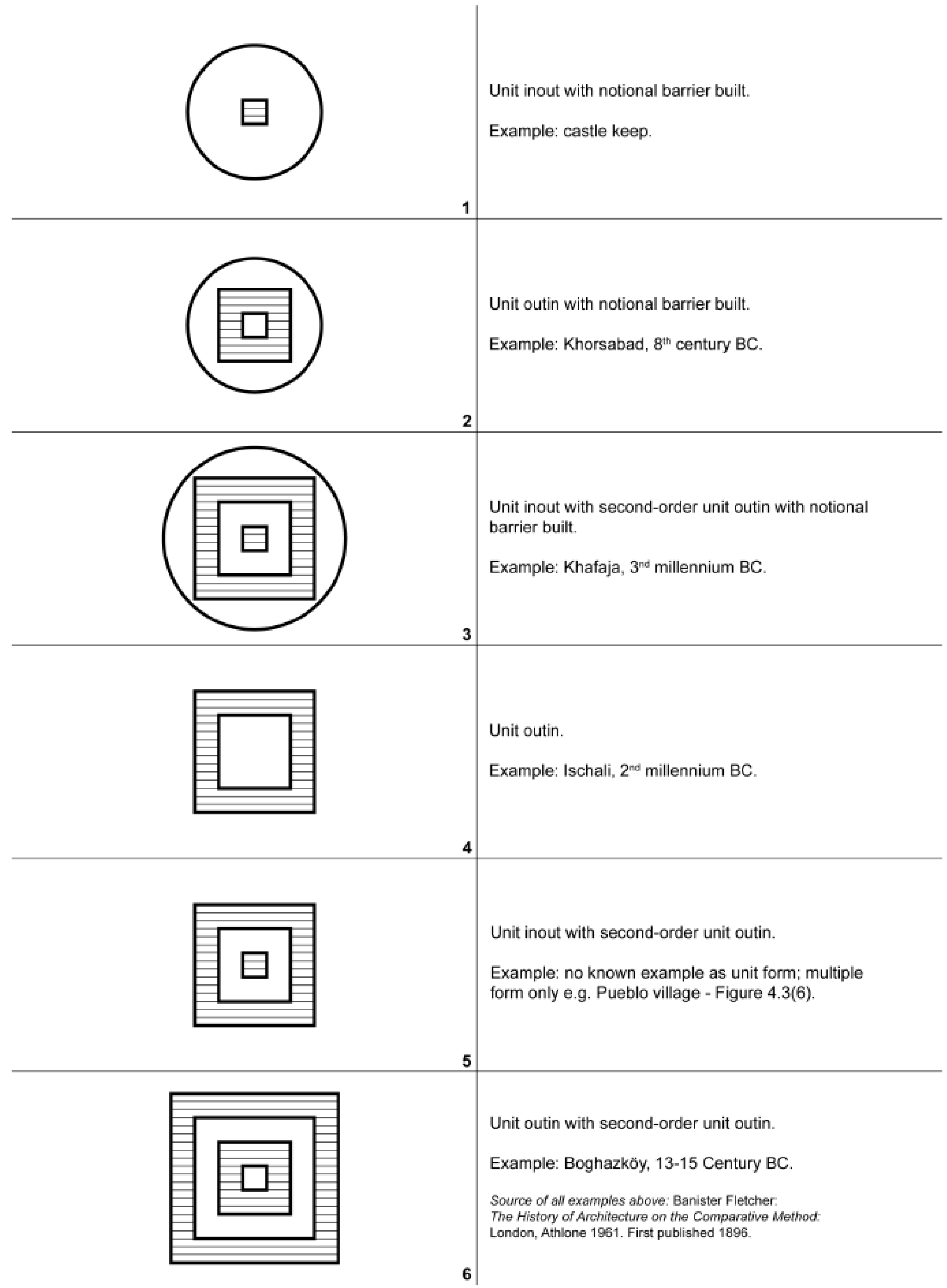


Figure 4.4 Examples of unit forms.

societies relatively independent of each other, and also for its structural stability until modern times. It also offers a formal set of reasons for the rapid transition from early settled life in villages to the first elaboration of proto-urban forms. Much work, however, remains to be done, and initial hypotheses are all that can be offered within this paper.

Semantics of spatial surfaces

The work of Piaget has revitalised the question of the relation between mathematics and life. By showing that structures of algebra and logic can serve as models of explanation for cognitive behaviour – he offers them literally as 'internal' models of intelligence – and that these may originate in the concrete operations of the infant (more precisely, in their coordination), Piaget has linked the foundations of mathematics to the foundations of behaviour and thought. Structuralist linguistics and anthropology suggest that this approach can also be applied nearer to the surface of perceived reality, and that mathematical structuralism may provide as fertile a form of explanation in the non-natural sciences as it does in the natural sciences. An underlying theme of this paper is that the many domain structures of categorical algebra may provide models for 'meaning' and intelligibility in artificial systems.

The discovery that mathematical apparatus used in advanced areas of natural science can also serve as internal models for the cognitive behaviour of children and the social arrangements of 'primitive' people (group theory for example is used both in theoretical physics and in the study of kinship systems) is not a surprising convergence if proper account is taken of the duality of any cognitive process, between the figurative properties of the universe as it presents itself to our perceptions, and the ordering operations which the perceiving subjects applies to that universe to give it intelligibility. Classification, for example, one of the simpler cognitive processes, exhibits this duality by requiring two mental gestures to complete an act of classification: first, an intensional act by which the defining property, or intension, of a class is named; then an act of enumeration in terms of that property by which the extension of the class is established. The intensional act refers to the operational ordering of phenomena in terms of some structured view of reality in the subject, the extensional act to the figurative properties of entities enumerated in terms of that ordering.

This duality shows that any cognitive process must be seen as a dialogue between the ordering operation of an active subject, and the morphological properties of the universe. 'Primitive' people, in bringing order and intelligibility into their universes, naturally make use of the logico-algebraic ordering apparatus of 'intelligence'. That they do so, and do so at the level of observables, and also that they construct their artificial signifying systems using the same apparatus, has been the theme of Lévi-Strauss. The discovery of logico-algebraic structures as ordering principles in primitive cultural systems is not surprising. They arise from the operational side of cognitive activity. However, such a simple ordering process, limited to observables, inevitably generates anomalies, which the elaborate strategies of myth attempt to explain away.

By contrast, science is intolerant of anomaly and looks beyond the inconsistent surface of things for internal ordering processes. But science also adopts a dual strategy, reflecting the dual nature of the cognitive act. It makes explicit its idea that there is 'significant connectedness' in the morphologies that are its object. In other words, science lays bare its theories in order to test and, if necessary, change them. Further, science makes explicit its own operations by constructing formal symbolic systems which embody the operational aspects of cognition. These are laid out besides the exposed theory so that tests of the mapping between the two can be made and may be explored in an externalistic way. In doing so science progressively discovers and elaborates the internal transformability of mathematical structures, and increasingly these provide models for the transformability of natural morphologies (and, like language, provide in themselves objects of interest for scientists to study).

This leads to the identification of two levels for the application of logico-mathematical structures. First, the *everyday level* of their *unconscious* and unreflective application in order to produce, out of the observable universe, intelligibility and significance. This results in the continuous construction of artificial systems of significance which we call 'cultures' which embody the transformability of logico-algebraic structures, discovered through cognitive action. Second, the *scientific level* of the *conscious* exploration of the transformability of logico-algebraic structures in relation to hypothesised internal ordering properties of morphologies. When we apply the second to an understanding of the first, as when anthropologists study primitive peoples, or psychologists experiment with children, we discover an unexpected isomorphism between mathematics, including its transformability, and real life. It is clear on closer inspection that this should not surprise us.

These considerations may help to understand what often appears as a basic paradox in the relation of logic and life: logico-algebraic structures appear to be of the nature of all artificial systems, yet to explain none fully. Logic, it appears, is too pure to represent the internal ordering of artificial systems like cultures. The requirement appears to be that imperfection should be taken account of in an ordered way.

But this apparent paradox only results from the covert assumption that the logico-algebraic structures of cognitive operations are actual properties of the universe to which the operations are applied. For example, to say with respect to any binary distinction that an object either has or has not the property, *Y*, is a statement about cognitive operations upon the world, not about the world itself. With respect to logical 'truth' the misunderstanding is more far-reaching. Logical truth is a production of the symbolic system of logic, and says nothing about the world to which it refers. To produce 'scientific truth' from 'logical truth' requires a perfect mapping between the domain of logical symbols and the real world, but this mapping would itself have to be represented in a perfectly mappable set of symbols, and this leads to an infinite regress.

With regard to the man-made semantic universes we call cultures, a parallel point can be made. Just as the logical operation of the constructive subject does not imply a parallel ordering of the real universe, so the logical method of constructing systems of signification does not mean that the resultant semantic field will itself exhibit such a simple ordering. It is much more likely to exhibit the result of a logico-algebraic procedure operating in a universe which consistently fails to respond to the simplicity of its operations. The initial question should be: how can logico-algebraic operational structures cope with the continual discovery of the imperfect results of its own perfect operations, and a universe that continually poses problems for our logical activity?

The history of 'thought about thought' is full of the awareness of such imperfection. The Chinese Yin-Yang symbol – a disc divided in an S-shape, but with each half of the division carrying a remnant of the other – expressed its dominance in Chinese thought. Leach has commented on the importance of anomalies in classificatory schemes in, for example, providing animal categories for swearing and other purposes.[13] In physics itself, the imperfection of a process came to be seen as its most important physical property. There are many other examples.

Boolean algebra (taken here as the simplest paradigm for logical operations), however, is founded on the exclusion of imperfection. This exclusion is expressed through the Boolean concept of 'negation', and the related concept of 'universe'. In any universe, $\cup$, with objects, x, then the property p or not $-p$ accounts for all objects in the universe:

$$[x: p] \cup [x: \neg p(x)] = \cup$$

and this is true for an extension of the universe. The structure of assumptions says that a universe exists (not necessarily a 'universal' universe, but perhaps a defined universe of discourse) and that the operation of negation in terms of any property divides that universe into two parts with no remainder.

In terms of the cognitive duality, it is clear what is happening. Boolean algebra provides a procedure for handling the extensional side of cognitive activity, taking for granted both the universe and the intensional structures representing this universe in the subject. If anything, it is a process for reordering the relationship between the two according to a simplified but more rigorous rule system. It provides a model for the establishment of pure and logical relationships but says nothing about either the universe to which it refers and takes for granted, or about the intensional structure of the subject which it also takes for granted as a source of defining properties for the extensional activity that it orders. The Boolean is an external model builder unlikely to recognise that his simplified, but purified, operations say nothing about the universe itself, nor about the intensional structure through which we order it.

However, in approaching the semantic fields of artificial systems, it is exactly the intensional aspects that are the object of interest. It is not clear what contribution the Boolean model could make to such a study; first because it is extensional and reformist rather than intensional and analytic; and secondly because it assumes the existence of a universe, when the semantic analysts must seek to understand how such a universe can come to be *constructed* through the logico-algebraic activities of individuals and societies. If he assumes the universe and models it, a Boolean externalist implicitly assumes that the logico-algebraic ordering structures which generate the universe are as simply represented in the universe itself. Such an approach invariably breaks down under an apparently unmanageable complexity.

It appears then that logic, as we have it, does not present useful models for the recapture of the internal logical order of the semantic fields through which the artificial universe is experienced. What appears to be required is a procedure which does not assume, but investigates, the relationship between a logico-algebraic ordering procedure, and a universe which resists its simplicity. The scientific model of the *dialogue* between the formal symbolic and morphological domains appears to be exactly what is required. The semantic fields resulting from such an enquiry would not simply be ordered in logico-algebraic terms but would be the resultant of a dialogue between logico-algebraic structure and a universe with infinite morphological variety. Since modern science characteristically represents a shift from an interest in extensional structure to an interest in intensional structure, such a procedure would be in keeping with the spirit of science.

The question may therefore be put directly: what would be the effect of running a logico-algebraic structure, as similar to a Boolean algebra as possible, in a universe which resisted its simple discriminations at every turn? The answer is, so far as can be seen, that a paradigm for the evolution of semantic universes and their structurally stable states is discovered. The following paragraphs attempt briefly to characterise such a structure, which appears to yield an account of the semantic evolution of spatial forms. Within the scope of this paper, formal considerations are dealt with briefly, since the mathematical and axiomatic aspects of the approach are dealt with in another paper in preparation, concerned solely with this subject.

The formal system proposed is called 'semantic algebra'. It has many of the structural properties of a Boolean algebra, although its outcomes are quite different and adjustments are made all the way through, following the requirements of changes in the axioms. For the purposes of the following explanation, the general structure of Boolean algebra is assumed as a starting point, although this assumption is not made in the axiomatic version of the system.

The perfection of Boolean algebra is embodied in the Boolean concept of negation and its linked concept of universe, which are expressed in the fundamental equations: $A \cap A' = \varnothing$ from which it follows that $A \cup A' = \cup$. Semantic algebra introduces a new *axiom of imperfection* $A \cap A' \neq \varnothing$ from which it follows that $A \cup A' \neq \cup$. In a semantic algebra no construction of A, A', $\cap$ and $\cup$ can constitute a finite universe, from which, it might be suggested that the 'universal universe', $\cup_{\cup}$, cannot exist: $\cup_{\cup} = \varnothing$. This gives the *axiom of non-existence,* which states no universe exists, nor anything in it exists, until it is made.

The axioms of imperfection and non-existence mean that any dichotomy implies that an object can be made which embodies *both* the positive and negative properties (the axiom of imperfection); and that a semantic universe is one that is constructed, *not* one that is found and subsequently ordered. In other words, nothing in a semantic universe exists until it is made, *not even the universe itself* (the axiom of non-existence).

More precisely, the concept of negation and universe have both been 'localised', to conform to the requirements of constructive and unfolding systems, as opposed to systems that already exist. In semantic algebra the term 'universe' is used to mean the set of objects constructed by the operations of semantic algebra, which means that a series of universes is created as the algebra 'unfolds'.

The localisation of the concept of negation is the foundation stone of semantic algebra. As Wittgenstein notes, in real life negation appears to *imply* what is negated, but 'not the *assertion* of it'[14] (our emphasis). In fact, three types of binary dichotomies recur throughout language. The first is what can be termed the 'positive inverse'. The 'positive inverse' of the verb 'give', for example, will be an equally definitive gesture in the 'opposite' direction, as realised in the verb 'take'. Secondly, there is a 'universal' or 'logical' negation of the verb 'give' which is simply expressed as 'not give', or in other words everything in the universe which is not 'give'. Boolean algebra depends on 'universal negation'. The third possibility is that the verb 'give' also has a negation which is inextricably linked with the term negated, but says nothing at all about the remainder of the universe. For the verb 'give' this 'local negation' is the verb 'keep'. 'Keep' *implies* the notion of 'give' but does not assert it. 'Keep' raises the possibility of 'give' only to deny it. The relation of 'keep' to 'give' is an example of the paradigm of *local negation,* which is the basic concept of negation used in semantic algebra. In the same way, the concept 'zero' is the local negation of 'number'.

The concept of 'local negation' raises a further possibility which the other forms of negation do not appear to possess. A concept may be combined with its local negation to construct a more complex concept. In relation to 'give' and 'keep', for example, the construction 'give keep' has one of its realisations in the verb 'lend'. This operation of construction resembles the Boolean operation of intersection, although with quite a different result. This homology can be clarified by means of a Venn diagram, as in Figure 4.5, which illustrates both the operations, 'local negation' and 'intersection', and the concept of a local universe. The construct 'givekeep' is said to be the universe of its members, 'give' and 'keep' and itself.

A semantic algebra is a procedure for unfolding a simple concept like 'give' into a complex semantic field. It consists of one object, namely a concept of logos, two operations, namely local negation and intersection (the combining of an object with a local negation), and a rule that states that all operations that are possible must be carried out. The result of this structure is an unfolding series of universes

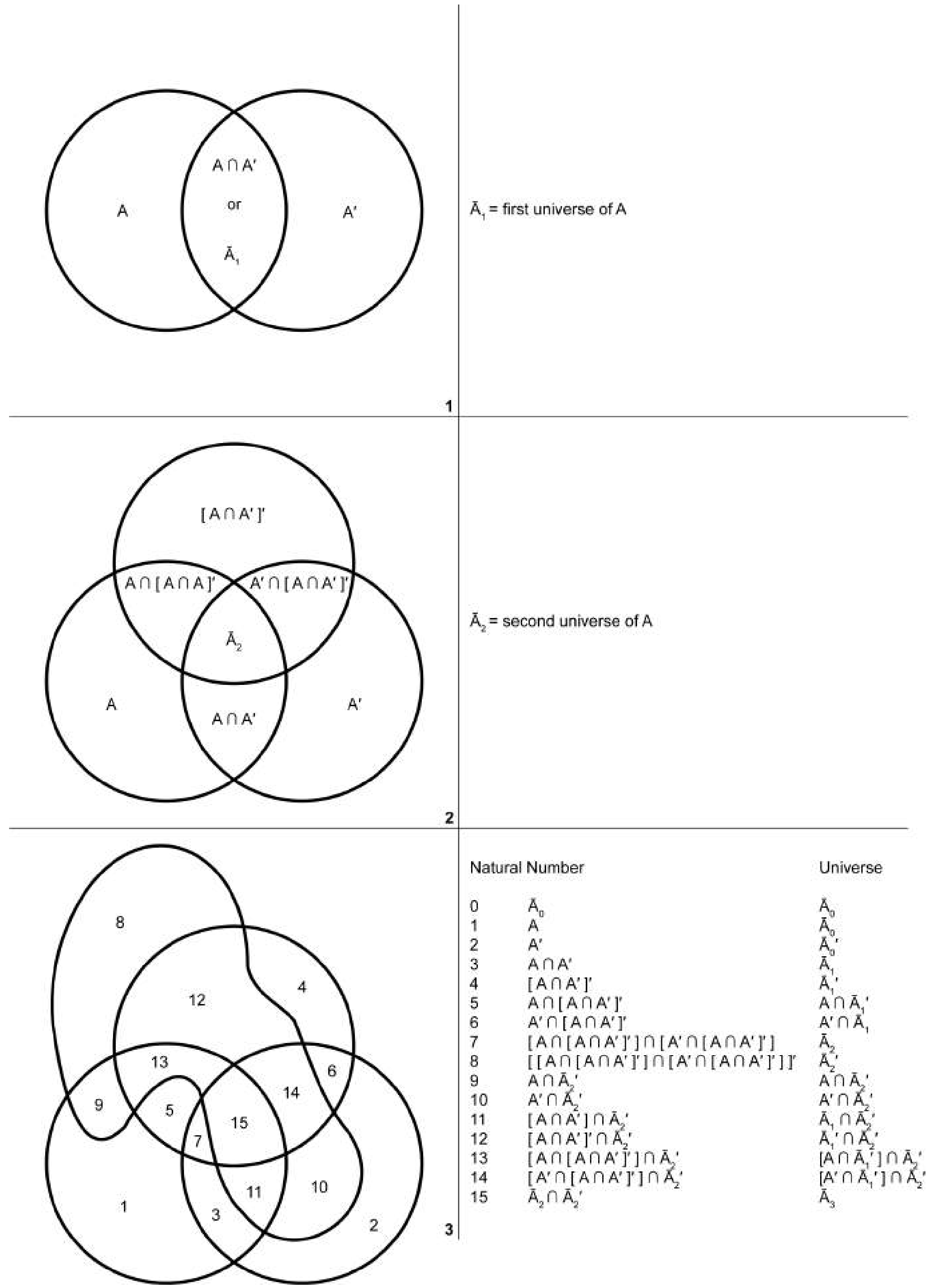


Figure 4.5 Unfolding of a semantic algebra (Venn diagrammatic and algebraic forms).

of more and more complex objects and structures, which are offered as an internal model for the evolution of complex social concepts out of simple individualised gestures and states.

Given the construction of the first universe shown in Figure 4.5$_1$, the operations are then applied again. Since it can be shown that double negation is not equal to affirmation in a semantic algebra, and that neither the initial object nor its local negation are part of the local negation of the first universe, the unfolding proceeds by creating a new object outside the existing structure (the local negation of the first universe), followed by the intersection of this structure with all previous structure with which it can intersect without tautology, that is without being reduced to an earlier structure by resorption (for example the intersection of A with $A \cap A'$ is $A \cap A'$ but the intersection of A with $[A \cap A']$ is the structure $A \cap [A \cap A']$. Figure 4.5$_2$ shows the construction of the second universe, and Figure 4.5$_3$ a third universe, though here numbering is used to show that the unfolding can be mapped one-to-one into the natural numbers and is therefore denumerable.

This unfolding structure possesses many properties that may not be obvious at first sight. First, any universal structure (identified by the heavy capitals, for example, $\overline{A}_1, \overline{A}_2, \overline{A}_k$) contains as part of itself all previous structures in the unfolding. This corresponds to semantic reality, in those terms ('village', 'town' and so on) represent exactly such complex, many-faceted concepts, hence the difficulties encountered in defining them indicatively. Also, through this property a link may be forged between semantic algebra and information theory. Second, notions like 'forgetting' and 'remembering' are representable, in that a local negation of a universe 'forgets' the internal structure of that universe, by treating it as a 'gestalt', but then 'remembers' previous structure by intersecting with it to make new structures. In the same sense, the act of 'design' forgets the internal structure of a multiple form, by treating it as a unit form, this internal structure then being remembered by active elaboration and usage. Third, the unfolding is modelled on the same principle as the spatial syntax outlined in the 'previous section: namely that the same operation may be applied to structures generated by operating on simple objects. to produce more complex structures. This provides a possible key to the retention of isomorphism between dissimilar domains as artificial universes evolve. Fourth, it allows the logico-algebraic construction of semantic fields which to the experiencer exhibit rich properties of ambiguity and denseness. Fifth, the graph representation (Figure 4.6) shows that a semantic algebra resembles the 'inverted pyramid hanging by its vertex,' suggested by Piaget as the proper intuitive model for human knowledge. Sixth, a semantic algebra provides a paradigm for context sensitivity in a meaning system in that any structure is intelligible as itself, and as a member of its local universe, paralleling the intuitive understanding of, say, being in a village and being in a particular space in a village.

The seventh property is more remarkable. The sequence of operations on the non-universal structures has the effect of reproducing exactly the same unfolding in the higher-order universes. In other words, a simple, algebraic relation between

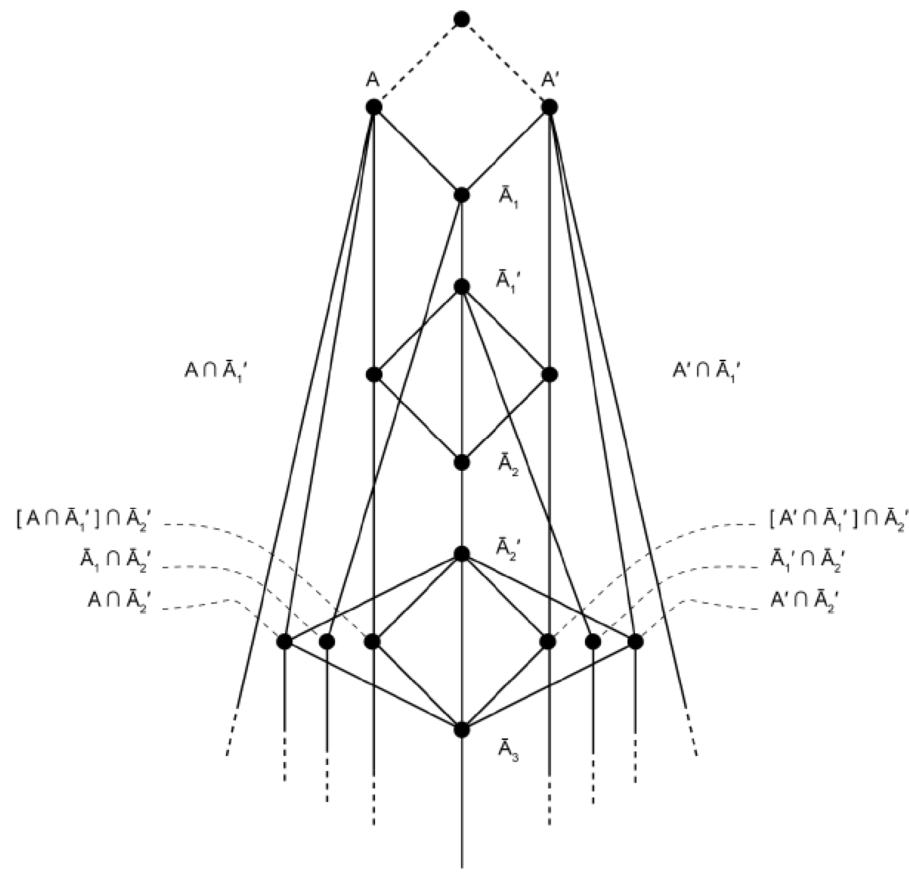


Figure 4.6 Unfolding of a semantic algebra to third universe (graph theoretic representation).

hierarchic levels appears to be generated. This is an interesting result, because understanding relations between levels in hierarchies is a leading problem in many scientific areas.[15]

Finally, semantic algebra shows how the concepts of structural stability and morphogenesis may exist and be ordered in a semantic field. This suggests a possible relationship to the work of René Thom on morphological archetypes in language,[16] since semantic algebra suggests a logico-algebraic regularity in the production of complex social concepts out of individualised primitives and offers a paradigm for relating Thom's exploration of the semantic basis of language to the complex semantic system in daily use. Of course, language itself remains largely at the level of the first, or at most the second, universe. More complex unfoldings are not those represented in words but in cultural systems. As far as architecture is concerned, the thesis is that artificial space represents one of the systems within which these more complex socialised concepts are represented, in particular the unfolding of the verb 'to be'. This gives it a certain primacy as a social language and suggests that architecture may be much more important to society than the social scientists, with their mechanistic metaphors, have allowed it to be. Artificial space is, literally, a language in which society says things which are too complex to be said in words.

A worked example of the unfolding of a complex social semantic field from a simple individual gesture is given in Table 4.1. This example is used because the second universe of 'give' is 'general ritualised exchange' which appears as a structurally stable regulator in most pre-trade societies. The reader's attention is drawn to Marcel Mauss's book *The gift* where such systems are examined in detail, in which the exchange of objects and services is used as a primary set of social signifiers, to regulate relations within and between societies. A theory of the ubiquity as well as the structural stability of such forms is offered by the semantic algebraic interpretation.

One further point of some importance is illustrated by this unfolding. In its simple forms the verb 'give' requires a simple, probably individual subject, the 'giver'. But as the concept unfolds into its higher-order structures, a higher-order subject is required, an 'inter- subjective subject'. This suggests that the tendency to personalise

Table 4.1 Unfolding of the verb 'give' to second universe

G_0		logos 'give'
	G	to give
	G'	to keep
G_1	$G \cap G'$	to keep by giving, that is, to keep the symbolic value of giving (lend is a modern version) to the generalised other
	$[G \cap G']'$	the generalised other's keeping of the symbolic value of giving, that is, effectively, to be required to give to the higher-order subject's keeping (taxation is a modern version); obligation
	$G \cap [G \cap G']'$	the act of giving to the higher-order subject; ceremonial giving, as it were, to the set of social reciprocities (debased modern version: the charity dinner) (old version: ceremonial gift exchange)
	$G' \cap [G \cap G']'$	to give to the dead, who embody the generalised other, the higher-order social subject, but who keep without reciprocation; sacrifice (modern version: flowers on graves)
G_2	$[G \cap [G \cap G']'] \cap [G' \cap [G \cap G']']$	generalised ritualistic exchange as a means of regulating social relations (modern version: buying drinks, sending cards and so on)

society, and to ascribe to it thought processes, while being undesirable and naïve, may not be entirely devoid of structural reference. In the language of artificial space such 'inter-subjective subjects' are identifiable, as the unfolding of the verb 'to be', in relation to its various syntaxes, creates structurally stable states at its universal stages of unfolding. Such stable states are recorded in language in 'gestalt' terms like 'dwelling', 'village', 'town', 'city' and so on. The general form of a semantic algebra is best illustrated by considering the unfolding of the verb 'to be', given in Table 4.2, in relation to the different meanings assigned to the word, or phrases incorporating it, in everyday language. The initial meaning is to be physically present in a place, to 'be here', to be present in real space. The local negation uses the idea of physical presence in real space, but negates it. Both initial terms therefore relate to those aspects of our existence by which we are *connected to an immediate environment*.

The next term and its local negation refers to those aspects of our existence by which we are *independent of immediate environment*, that is, the world of internal representations and symbols. This might be thought of as the world of 'logical space' as opposed to real space. Everyday language uses the verb 'to be' in relation to logical space through such sentences as 'I am in the Communist Party', by which we say we are 'present' in an abstraction.

The following term and its local negation are constructed from the intersection of the two foregoing, in that *the world of representations and symbols is also constructed into real space*. It is possible to enter physically certain spaces which constitute real space realisations of logical space. Architecture participates in the realisation of such spaces. Its central theme is the mapping from logical into real space. The universe, A_2, is the structure comprising all three terms and their local negations.

Semantic algebra provides a model for the unfolding of all four domains of the code, namely logical differentiation, rules for elaboration, enclosure and permeability. All four are given up to the second universe in Tables 4.3, 4.4 and 4.5. Socio-spatial theory, it could be conjectured, arises from the dialectical relationship of the unfoldings of A and E.

Elementary relations in the AFPE code

The code properties of a spatial surface – differentiation, barriers, permeability and rules for elaboration – can be called 'logical properties', intelligible through understanding rather than through simple perception. The properties of any particular spatial surface result from an unfolding of E and P, through F which is given by A. F 'carries' the structure of A from the level of the simplest structure to the level of the spatial surface. Alternatively, it can be said that A is logical space, E is real space, and A is mapped into E through F and P. This may be represented by $A \xrightarrow{FP} E$ (hence the acronym *AFPE* code). Either way F, and its function, f, are basic to the decoding of a spatial surface, that is finding its generative formula.

Table 4.2 Unfolding of the verb 'be' to second universe: logical interpretation and spatial interpretation: 'A'

A_0		logos 'be'	spatial interpretation
	A	be present	places for people to be
	A′	be absent	places for things to be
A_1	A ∩ A′	be in domain of representations, symbols and so on, to be 'in your head'	places where people can be in their heads, for example, asleep, a garden etc.
	[A ∩ A′]′	for there to be a space of representations in which you are not, that is, in society's domain of representations	a sacred space where you are not in normal time, but where sacred things are
	A ∩ [A ∩ A′]′	to be physically in a space of society's representations and symbols	a ceremonial and profane activity space in which you both are physically and in society's domain of symbols
	A′ ∩ [A ∩ A′]′	not to be physically, but to be in society's domain of representations only, that is, to be a name only, to be dead	a burial ground in which you are not, in which you become a thing with a name
A_2		a simple society	a village

Table 4.3 Unfolding of 'enclose' 'E'

E_0		logos 'E'
	E	an enclosed space
	E′	a non-enclosed or defined space; an external space defined by the act of enclosure
E_1	E ∩ E′	a transition space, both enclosed and defined (for example, a garden or yard)
	[E ∩ E′]′	a social space set apart, and symbolically closed in normal time (for example, a sacred space)
	E ∩ [E ∩ E′]′	a space set apart yet accessible and symbolically open in normal time (for example, a public open space)
	E′ ∩ [E ∩ E′]′	a non-space set apart as a marker of defined space only (for example, a monument)

Table 4.4 Unfolding of 'permeable' 'P'

P_0		logos 'P' – 'permeable'
	P	permeable, having a hole to pass through (for example, an ordinary door)
	P′	not permeable (for example, a continuous wall)
P_1	P ∩ P′	permeable, but not to go through (for example, a window)
	[P ∩ P′]′	a symbolic barrier (for example, a church door)
	P ∩ [P ∩ P′]′	an open gate, symbolic entrance permeable in normal time (for example, entrance to public open spaces)
	P′ ∩ [P ∩ P′]′	the sealed entrance (for example, entrance to a tomb)

Table 4.5 Unfolding of rules for enclosure 'F'

F_0		logos 'F'
	F	free to enclose again
	F′	not free to enclose again
F_1	F ∩ F′	free to elaborate symbolically (grow flowers, place gnomes and so on)
	[F ∩ F′]′	not free to elaborate symbolically (since society's symbols are already present, for example, statues)
	F ∩ [F ∩ F′]′	free to enclose again, but only with a transient structure for profane public purposes which must be dismantled after use (for example, a place to put up stalls, fairgrounds and so on (a village green))
	F′ ∩ [F ∩ F′]′	totally preserve as it is (for example, the space around central public buildings, monuments and so on)

At a schematic level, the variability of spatial surfaces is given by relatively few relationships among the four domains, in interaction with certain social variables.

Call a random process a 'mindless' syntactic generation of a spatial surface out of *E*, unconstrained by any *E*, and by implication ignoring *A* (*P* being unimportant at this stage of the argument). If such a random process is allowed to generate an artificial spatial surface on a previously undifferentiated surface, the value of the function *f* of *F* will be seen to be maximal as the space approaches maximal denseness. In other words, a random process without a value ascribed to f is seen in time to be equivalent to f_4, or f_{max}. *It follows that insofar as the value of f remains less than maximal for a dense and stable spatial surface, then some structurally stable* *mechanism has intervened.* The maintenance of a structure of non-built or open space is therefore an indicator of exogenous ordering of some kind.

This leads to a distinction between two forms of density: semantic density and syntactic density. Semantic density exists when a spatial surface has been fully elaborated according to its generative formula. This will imply a greater or lesser degree of syntactic density, that is the degree to which the space is physically dense in that no further structure may be physically placed on the surface. By definition, an *f*-minimal surface will approach semantic density long before syntactic density, whereas an *f*-maximal surface will arrive at semantic density at a point much closer to syntactic density. An important result of this is that an *f*-maximal surface can therefore embody much more semantic content than an *f*-minimal surface, which arrives at density at a much earlier stage of elaboration. This is subject to the real-life constraint that if f is f_4, and the surface is dense, then no spatial surface exists in an experiential sense, only a series of discrete interiors. Thus, there is always some exogenous ordering of spatial surfaces.

Two fundamentally dissimilar lines of development for spatial surfaces arise from this distinction: those in which the value of f is minimal, in which case there is relatively strong exogenous ordering; and those in which the value of f is maximal, in which case the exogenous ordering is relatively weak, and the ordering of the spatial surface is more internal and syntactic. A natural relation with social variables follows. Insofar as a society uses space as a set of signifiers of social order, it will do so in terms of some exogenous ordering of the spatial surface in which values of f will be lowered or minimised; insofar as space is freed from the constraint of acting as a signifier for social order, its aggregates will tend to be syntactic and *f*-maximal. Thus, the use of space for representing social order tends to minimise f, whereas a more instrumental or free elaboration of space tends to maximise f. The invariance of this relation in the code is shown by the tendency, even in the densest urban spatial surfaces, for the main buildings embodying the social order to be surrounded by an open space barrier. The general differences in the spatial surfaces between the City of London, where only two buildings have open space barriers, and the City of Westminster, where many buildings have open space barriers, is given by this simple formula. The City of London reaches semantic density at a point much closer to syntactic density than the City of Westminster.

The difference between *f*-minimal and *f*-maximal spatial surfaces is the simplest and most general example of differential aggregation modes. It is best illustrated, as are the relations to underlying social variables, by the contrast between the spatial surfaces of pre-settled and settled societies, extensive documentation for which now exists in the archaeological and anthropological record.

Pre-settled societies, as studied by anthropologists, normally utilise space as one of the signifying systems for that form of social solidarity characterised as 'mechanical' by Durkheim, that is, a social solidarity dependent on symbol and ritual rather than on functional interdependence through extensive division of labour.[17] In general, such societies impose a strong exogenous and symbolic order on space, leading to *f*-minimal spatial surfaces at first and second-order levels.

This ordering begins to change with the advent of settlement. Settlement (as is too rarely observed by those who discuss the relation of social and spatial process) is a spatial act, meaning the fixing of work to place and the transition from a discoverable and untransformed local environment to one artificial and transformed. Settlement means that social space ceases to be primarily symbolic and becomes increasingly instrumental. The formative influences on settled space derive not from the symbolic systems of social order but from instrumental systems of survival and work. Lacking strong exogenous ordering of space, the spatial forms and the aggregate surfaces of the maximal values of f are progressively discovered and elaborated.

Since settled societies bring with them an inheritance of social complexity, the semantic content of the spatial surface does not change, but finds an alternative syntactic expression. At Çatal Hüyük,[18] the earliest known urban spatial surface, which is syntactically *f*-maximal, A is unfolded entirely in interiors and no use is made of the higher-order spaces generated by *f*-maximal unfoldings. The whole of the second universe of A, including its $\grave{A}_2$ form, the shrine, is unfolded at Çatal Hüyük through differentiation of interior spaces, urban form proper develops from the discovery and elaboration of the higher-order spaces of *f*-maximal syntaxes, which then become the bases for higher-order developments. The building block of the urban spatial surface is the *f*-maximal elementary form (the multiple outin form), which is also the basic form of an agricultural settlement, and the normal form for farms throughout the world. This, however, gives only one side of the formula for the urban spatial surface. Following Weber's definition of the city as a fortress/market,[19] a parallel duality exists in the spatial form. Insofar as the city is a fortress, concerned with external and internal social order, then the spatial outcomes are *f*-minimal at all levels. Insofar as the city is a market, concerned with productive work and exchange, then the spatial surface is *f*-maximal. This socio-spatial duality is universal in classical urban forms, invariant under technology, style and decoration.

Various other simple relations of this kind appear to hold. For example, matrilineal societies tend to the *f*-maximal first- and second-order (multiple outin) forms, whereas patrilineal societies use *f*-minimal first- and second-order (multiple inout) forms. Why this should be so may be clarified by consideration of the duality in the simplest spatial gesture of making a barrier, that is the duality of E and E'. An enclosure can either be a gesture of inclusion, E, or a gesture of exclusion, E'. The inclusive gesture faces 'inward' and does not preserve an open space beyond the barrier. The exclusive gesture, on the other hand, faces 'outwards' and defines an open space barrier beyond the indicated barrier. In a primitive military encampment, for example, this gesture is made twice: a gesture of exclusion, preserving an open space barrier, is made around the central building where spoil is kept; and then the exclusion gesture is repeated at the second-order level of encampment boundary, where an open space barrier is again preserved. It is noteworthy that the exclusion gestures refer initially to things, and the inclusive gesture to people, and this relation holds throughout the unfolding. A monument, for example, is not a simple spatial object but a highly complex one, defining only an open space barrier, and having no interior. The association of the exclusion gesture with the male role in society and the inclusion gesture with the female provides a rich field of study from the point of view of the battle of the sexes, seen as the earliest form of struggle between classes.[20] Town form is in effect a highly evolved synthesis of these contradictory lines of social and spatial development.

Urban form proper can virtually be defined in terms of an absence of open space barriers which have been, as it were, removed and placed on the outside of the town. It arises only after the exclusion gesture A_4, the building of a town wall, creates the conditions in which previous internal exclusion structure can be forgotten. At this point, the central citadel turns into a set of free-standing public buildings in a continuous landscape, and the internal court spaces of multiple outin aggregates open outwards and become side streets. The coming alive of the streets and the continuity of the elaboration of space through universes up to the fifth are the distinguishing marks of the urban spatial surface proper, as contrasted to the spatial surfaces of proto-urban forms, as occurred for example in pre-Columbian America and in early Egypt.[21]

The difference between an exogenously ordered *f*-minimal spatial surface and one internally ordered and *f*-maximal leads to another major dimension of the social meaning of artificial space; again so far as can be seen, an invariant in all cultures: namely the difference between sacred and profane space. A sacred space is highly structured in an intentional sense; it is fixed and finite, its 'meaning' is in its structure and it is exogenously defined. *A* profane space, on the other hand, has extensional variety rather than intensional structure; it is non-fixed and proliferates new aspects and forms of itself; its 'meaning' is in its variety rather than its structure and its structure is richly syntactic, rather than exogenously ordered. It uses syntactic possibilities to proliferate semantic content rather than using semantics to order syntax. Structure and variety are of course indissolubly connected in any meaning system, but one may be maximised at the expense of the other. Both are sought after, for example by tourists; in fact, the two poles perfectly characterise the type of spatial structures frequented by tourists: the places of maximal intensional structure, and the places of maximal extensional variety.

It is not possible within the scope of this paper to offer a complete account for the evolution and code form of the urban spatial surface, but a few notes on some of its more interesting logical properties are pertinent. First it is a *combined* unfolding maximising *f* for the most part, minimising it for certain areas. The *f*-maximal areas are subject to: (a) a high degree of permeability; linked to (b) the maximal elaboration of marginal space (that is space left over by enclosures at first- and second-order levels); and (c) subject to the structure given in Figure 4.7, which represents a simple linking of two semantic algebras, producing both 'normal'

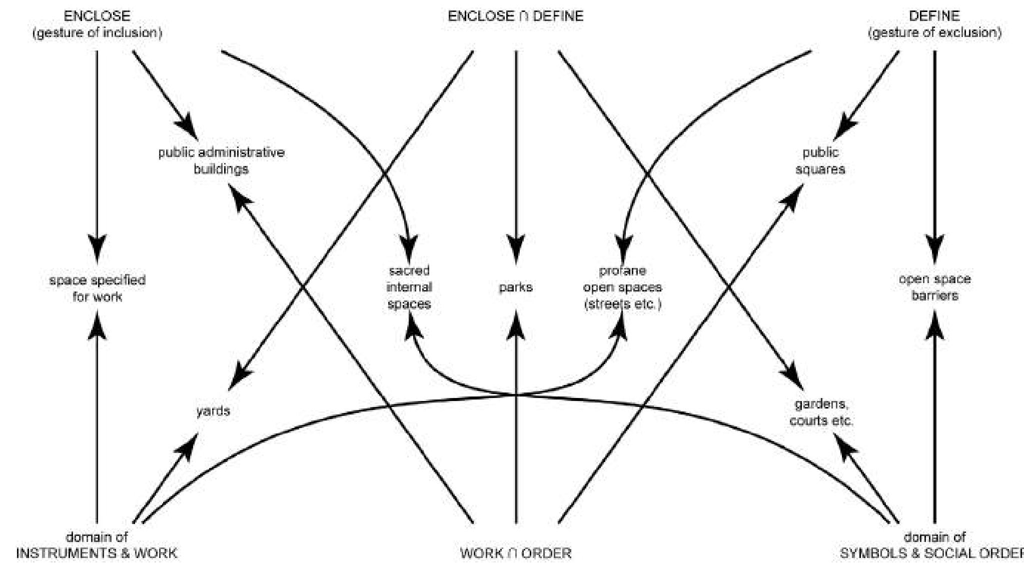


Figure 4.7 A semantic field, giving simple linking of two semantic algebras.

forms (verticals and horizontals) and 'special' forms (crossovers), the 'crossovers' being the forms of space that most characterise 'urbaneness' in the sense that the tourist will seek them out.

Of particular interest in this model are the 'marginal' spaces, that is, the 'by-product' external spaces resulting from building specific space for relatively tied uses. Marginal space is generated continuously by *f*-maximal aggregation modes, and if no inhibiting rule exists, such space 'tends to use', and secondary elaborations of the space appear. The elaboration of marginal space is one of the keys to the richness of the urban spatial surface in the profane areas of the town, and reconciles multiple use patterns with the natural predominance of space that is specified for a particular use. Also interesting are the 'ambiguous' spaces, since their nature is to be subject to the rule 'both build and not build' – that is build, but only non-permanent structures, for example, mobile markets and fairs. This ambiguous rule is a good demonstration of the importance of ambiguity in the elaboration of spatial surfaces, and how ambiguity can nevertheless be generated – and recaptured – in a structured way.

The logical nature of the modern transformation of urban space can also be clarified a little. The shift in general is from (a) the outin aggregation forms to the inout forms; leading to (b) the proliferation of pavilion forms in an open landscape surrounded by open space barriers; (c) the consequent freezing of most marginal space and secondary elaboration; (d) the establishment of clear, unambiguous demarcations and boundaries, where previously a structured ambiguity prevailed; and (e) loss of continuity in the surface. In general, these changes could theoretically result from the simple instruction in the code: prime all terms, that is make all gestures of exclusion. The new open spaces of the city are not made free by this transformation, but on the contrary subject to a strong exogenous order.

Such an explanation of the changing logic of the urban spatial surface would appear to be explicable through the given relations to social variables. It could be argued that the urban spatial surface was the spatial equivalent of 'organic solidarity', that is, a society dependent for its internal structure not so much on symbols as on functional interdependence induced by the extensive division of labour. Organic solidarity requires the functional integration of space, and therefore all f values are maximal wherever possible in the urban spatial surface, subject to the existence of a ruling authority whose structures are *f*-minimal. It appears that in the current transformation, space tends to be *f*-minimal rather than *f*-maximal, and it can therefore be hypothesised that a structurally stable exogenous order is being imposed, which represents a concern for social order rather than for work. A general impoverishment and fixity would result from this, and a greater emphasis on aesthetics combined with a dislike of non-formal elaboration of space. All these properties appear to be present in current spatial orthodoxies, especially those whose prime concern is with aesthetics.

An explanation for this could be that the type of division of labour that exists today is dissimilar in type to that of the urban organic solidarity. In the phase of industrialism, people do not create patterns of functional interdependence by making relations with each other, but instead each individual makes an arrangement with a focal point. A series of pyramids rather than a network is the result. The emphasis is on the individual, with a unique relation to his focal point, but lacking lateral relations with his fellow men of a direct and functional kind. In such a situation, a society might be expected to return more and more to conditions of mechanical solidarity, and use symbolic means more and more to stabilise the social order and yet confirm the separation of members from each other. This would link the primary role of media in our type of society with the transformation of space into what it was prior to large-scale settlement, namely a signifying system presenting the social order in a society dependent on such mechanical solidarity devices for its cohesion. In such a situation it is difficult to see how architecture could function other than as a device for covert social control. Indeed, this has been the primary theme in the social discussion of spatial form for the past one hundred or more years.

The theme of the double boundary is also important to the modern transformation. The *f*-maximal outin aggregation mode generates double boundary structures from the inclusion gesture, giving buildings surrounding a space in a many-to-one relation, whereas the *f*-minimal pavilion forms currently in use, with their open space barriers, generate double boundary structures from the exclusion gesture. In the latter case, open space becomes a no-man's-land between the inhabitants and the outside world, rather than a communal interior space, permeable to the outside yet unique to the inhabitants. In semantic algebraic terms, the open areas surrounding a block of flats are $[E \cap E']'$ (exclusion), whereas the street was $E \cap E'$ (inclusion form).

These changes in terms of open space barriers, double boundary structures (and multiple boundary structures of the urban spatial surface), marginal space and its elaboration, permeability, strength of boundaries and classifications, show how profound the change is in the logic of space in our time, and how fundamental its discontinuity from spatial history since the beginning of human settlement. These changes are related to social changes, and the sources must be sought in the debate on the relation of space forms and the maintenance of order. In this situation, the synthesis of space as a means of achieving social order with the architectural imagination, as manifested in the work of Le Corbusier and others, might be thought inimical to the evolution of spatial surfaces that reproduce the richness and social viability of earlier structurally stable forms.

However, it is possible that this present situation, in which the new logic of space arises from the generalisation of the exclusion gesture, and the negative aspects of spatial evolution, will be superseded by changes which are now beginning. Most probable is that the urban spatial surface, which was abandoned with the town itself, will return (or will be 'remembered'), but without finite urban forms which led us to think that building outside their limits had to follow extra-urban aggregation modes. The result will be not the universal garden city, but the discovery of new and richer forms of the urban spatial surface. The evolution of such forms of artificial space will depend on understanding the internal generative structures of a non-imposed spatial order, in which the complex of rules for its manufacture will be distributed in the society rather than located in a single focus, and in which the role of designer will be the generation of structurally stable unfoldings rather than the imposition of exogenous order.

Acknowledgements

The authors wish to acknowledge the help of the following: Professor Pat O'-Sullivan of UWIST (University of Wales Institute of Science and Technology) for making key suggestions that led to the development of semantic algebra; Professor John Musgrove of UCL School of Environmental Studies for providing an academic milieu in which this work could advance; the MSc students of 1972/3 of the School of Environmental Studies at UCL to whom we owe too many debts for individual acknowledgement, and in particular to Mick Bedford, who drew the diagrams and criticised the text.

Postscript

Adrian Leaman

It was supposed to be simple.

Build a wall, what happens? A wall modifies climate; displaces resources; is a symbol of some sort; and modifies behaviour. A wall does all of these things whether you intend it or not, like it or not.

Then you need a knowing subject. As I sit here, I'm well aware of the inside and outside, and the window between. The rest of the world is also out there, but not so obvious or relevant to me at this moment. The reason that I'm inside now is that the arrangement of walls and windows working as a stable relationship offers more potential to achieve a greater number of useful activities, like writing this.

Even in one paragraph it is possible to grasp what needs to be covered in any description of how buildings and settlements come to be as they are. The notation will include walls, patterns, rules about their arrangement, enclosure, openness, inside, outside, the natural world, the environment, the artificial, the knowing subject, symbolism, boundaries, categories, transitions, movement and, of course, space and time.

Working from our office in the RIBA in the early 1970s, Bill Hillier and I were often distracted from what we were supposed to be doing by what we really wanted to do: namely, come up with a theory which would help explain the richness, variety and meaning in human settlement patterns. We looked to spatial analysis, linguistics, logic and anthropology amongst others: anywhere but the received wisdom of the existing environmental disciplines. Our territory was the 'Sciences of the Artificial', as Herbert Simon named them: concepts like chaos, adaptive systems, complexity, uncertainty, epigenesis and morphogenesis. We had no access to computers. This was still the era of punched cards and single multiple regressions that took a week to compute and check by hand.

The basic metaphors for our efforts became syntax and semantics, then genotypes and phenotypes. Not ideal because they are derivative, but good enough to help us explore likely codes and their social manifestations. Unfortunately, the handle 'space syntax' hijacked our efforts generically, and the semantics side was increasingly sidelined.

As John Peponis has said so admirably above, the semantic algebra proposals were never followed up and concepts like 'local negation' left hanging for others. Hamlet was not quite right. It's not 'To be or not to be'. You need to substitute 'and' for 'or' and add some brackets to put it like this '((To be) and (Not to be))'. A space semantics built this way will say that you can have a real construct like 'Not((To be) and (Not to be))'. You will find it in plain sight everywhere.

September 2022

Notes

1. R. Thom, 'Comment by René Thom', in *Towards a theoretical biology*, edited by C. H. Waddington (Edinburgh: Edinburgh University Press, 1968), pp. 32–41, at p. 32.
2. Quoted in E. Fromm, 'The significance of the theory of Mother Right for today', in *The crisis of psychoanalysis* (London: Jonathan Cape, 1971), pp. 100–105.
3. B. Hillier and A. Leaman, 'How is design possible?', *Journal of Architectural Research* 3 (1974): 4–11.
4. R. Thom, 'Topological models in biology', in *Towards a theoretical biology*, edited by C. H. Waddington (Edinburgh: Edinburgh University Press, 1970), pp. 89–116.
5. Hillier and Leaman, 'How is design possible?'.
6. See, for example, M. Echenique, 'Models: A discussion', *Architectural Research and Teaching* 1 (1970): 25–30.
7. B. Hillier and A. Leaman, 'System, structure and transformation', *Transactions of the Bartlett Society* 9 (1973): 36–77.
8. T. Winograd, *Understanding natural languages* (Cambridge: Cambridge University Press, 1972).
9. Editors' note: Hillier and Leaman may be referring to L. March, 'Models, modes and mores', in *LUBFS conference proceedings: Urban development models* (Lancaster: Construction Press, 1975), pp. 301–310.
10. Editors' note: Hillier and Leaman are likely referring to the Centre for Land Use and Built Form Studies founded by Lionel March in the late 1960s at Cambridge in the School of Architecture, where he and colleagues were exploring built form mathematically.
11. B. Rudovsky, *Architecture without architects* (London: Academy Editions, 1964); D. Fraser, *Village planning in the primitive world* (London: Studio Vista, 1968).
12. B. Fletcher, ed., *A history of architecture on the comparative method*, 17th edn, revised by R. A. Cordingley (London: Athlone Press, 1961) (1st edn 1896).
13. E. Leach, 'Anthropological aspects of language: Animal categories and verbal abuse', in *Mythology: Selected readings*, edited by P. Maranda (London: Penguin, 1972), pp. 39–68.
14. L. Wittgenstein, *Tractatus logico-philosophicus* (London: Kegan Paul, 1922), p. 75, note 4.0641; and L. Wittgenstein, *Philosophical investigations* (Oxford: Blackwell, 1953), para. 447.
15. H. H. Patee, 'The problem of biological hierarchy', in *Towards a theoretical biology*, edited by C. H. Waddington (Edinburgh: Edinburgh University Press, 1970), pp. 117–136.
16. R. Thom, 'Topologie et linguistique', in *Essays on topology and related topics* (Berlin: Springer, 1970), pp. 226–247.
17. É. Durkheim, *The division of labour in society*, translated by George Simpson (New York: The Free Press, 1933) (1964 English edn).
18. J. Mellart, *Çatal Hüyük: A neolithic town in Anatolia* (London: Thames and Hudson, 1967).
19. M. Weber, *The city* (New York: The Free Press, 1921) (1968 English edn).
20. For example, Mellart, *Çatal Hüyük*.
21. E. Fromm, 'The significance of the theory of Mother Right for today'.

5 Architecture as a discipline (1976)

Introduction to 'Architecture as a discipline'

Frederico de Holanda

From the beginning of the 1970s, Bill Hillier wrote a series of articles co-authored by Adrian Leaman. Of those texts I would select 'A new approach to architectural research' (December 1972), 'How is design possible?' (January 1974), and 'Architecture as a discipline' (March 1976) as the most emblematic – we might call them the 'Holy Trinity' of the theoretical foundations of what would be named *space syntax*, an expression which appeared in print form only in December 1976.[1] The axioms, concepts, method and techniques of the theory would be fully exposed eight years later, when Hillier co-authored with Julienne Hanson their fundamental book *The social logic of space*.[2]

The three articles have in common such fundamental ideas as *meeting ground theory*, architectural *functions*, architectural *codes*, *means* and *ends* research, and *environmental sciences*, presented in each new article in new facets. As follows, an inevitably personal view informs the selection and relative importance given to some of those concepts.

Perhaps the key aspect is to confer on architecture the status of a *scientific discipline*. The 'dominant research question' turns to the 'what' instead of the 'how', the emphasis that had hitherto characterised research: on *practice* and on the *nature of design*, not on the *architectural object*. The view of a 'meeting ground' field of thinking, by which architecture would simply constitute an instance of *application* of knowledge produced elsewhere – that is, in the consolidated social sciences – should be replaced by reflecting on how buildings, neighbourhoods and cities perform in response to human needs. (This hegemonic view pervades not only academia but also governmental research funding agencies such as the Brazilian CNPq, where architecture is classified as 'applied social science', in contrast with the 'human sciences' of sociology, anthropology, economics and so on.)

Hillier and Leaman submit the concept of *function* as architectural performance at all scales affecting people's life. A 'four-function model' suggests architecture works as a set of *modifiers* of pre-existing reality. Labels vary between the three papers, but they are consolidated, in 1976, as *climate*, *activity-space*, *economic* and *social language* functions. In other words (which are *not* in their text), Hillier and Leaman suggest that the necessary shift is also one from *architecture as a dependent variable* to *architecture as an independent variable*. To rest on consolidated disciplines implied that analytical categories belonged to *them*, and that architecture was simply viewed as the *outcome* of economic, social, political, (natural), environmental and other issues (therefore, *dependent* variable). What now sounds like a truism – that buildings and cities *affect* us according to their attributes – was seen before as merely epiphenomenal.

This 'interdisciplinary fever' swept architectural schools at the time around the world, the UK and Brazil included, in research and in design. I entered academia in 1972 at the Faculty of Architecture, University of Brasília, the year of the first text referred to above. Having come from professional architectural practice, I was surprised to see that it was difficult to convince students to embark on . . . architectural design. A 'serious' work had to be produced in sociological, economic and political jargon, and had to deal with their correlate analytical categories – anything but architecture. And – alas! – to produce architecture was considered 'rightist'. Bar rare exceptions, the result was shockingly mediocre.

And yet, there is nothing *against* interdisciplinary contributions to architectural knowledge and practice in Hillier and Leaman's texts, or against teamwork across academic borders; on the contrary. However, bridges towards established disciplines must be launched *from* architecture, the description of which varies according to knowledge objectives. The focus is on architectural functions, and the potential contributions from sociology to 'activity-space function', or from semiotics and aesthetics to 'social language function', or from economy to 'economic function' are all too obvious. But the beauty of the challenge is that, in this radically different 'meeting ground' stance, categories are not ready at hand – they have to be collectively constructed along those two-way thoroughfares.

Hillier and Leaman's reasoning travels far back in time. The idea of *architectural function* had a noble pioneer in Marcus Vitruvius Pollio (circa 27 BCE).[3] Vitruvius's writings are interesting because he deals with the 'meeting ground' idea, but he does so in both senses presented above. Architects should be acquainted with knowledge on clockworks, medicine, astronomy . . . But they should also care about those 'bridges' through a 'three-function model': *commoditas* (Hillier and Leaman's *activity-space*), *firmitas* (the building should not fall upon our heads) and *venustas* (Hillier and Leaman's *social language*).

If Hillier and Leaman's reasoning is deeply rooted in the past, it also challenges us to move forward, and to devise subdisciplines to approach the various functions of the model, as knowledge grows and specialisations bloom – for example, climate function into acoustics, thermal comfort and luminousness issues (in fact, these three topics are the subject matter of three different courses at our faculty). Another example further illustrates this: a 'research division of labour' among our staff in Brasília led to an 'eight-function model' that unfolds Hillier and Leaman's scheme particularly with regard to 'social language': topoceptive (memory and navigability), symbolic (representations), affective (psychological states) and aesthetic (sensory pleasure and world view) aspects of architectural performance are explored. This is likewise possible in other functions.

From the 1950s onwards, Modernism in architecture was strongly – and rightly – criticised, particularly at the urban scale: the renowned Jane Jacobs comes immediately to mind.[4] But the theoretical tools for such criticism had scarce impact on buildings and settlements for a simple reason: they were not *morphological*. It took another 20 years for architecture to be brought to frontstage in scientific reasoning, and this happened simultaneously in various countries – France (Philippe Panerai),[5] USA (Stanford Anderson),[6] Brazil (Carlos Nelson Ferreira dos Santos)[7] and, the focus of our attention here, the UK, with Bill Hillier as the leading figure of a team of scholars at the then Bartlett School of Environmental Studies, University College London (UCL), and as an outstanding thinker in the field of architecture worldwide. He has illuminated our academic and professional careers ever since.

Architecture as a discipline

Bill Hillier and Adrian Leaman

Originally published: Hillier, B. and Leaman, A. 1976. 'Architecture as a discipline'. *Journal of Architectural Research*, 5, 28–32.

Is the study of the built environment a subject in its own right, or is it simply the meeting ground for a number of disciplines? Should environmental studies be a loose faculty arrangement in the university with architecture as one of a number of 'related disciplines' grouped around a problem area? Or is there some sense in which the study of the built environment can arise naturally from the activity of architecture? The aim of this paper is to sketch a view of architecture as a discipline which looks into the nature of architecture itself for the disciplines and theory on which the academic study of the built environment is based.

The paper argues that it is possible to organise an approach to architectural and environmental problems based on the requirements of designers rather than on the academic structure we have inherited. This shifts the focus of research from the methodology of design to the nature of the building itself, while also making the connections between science and design much less of a problem.

Is architecture a discipline?

Is the study of the built environment a subject in its own right or is it simply the 'meeting ground for a number of disciplines'? Should 'environmental studies' be a loose faculty arrangement in the university, with architecture as one of a number of 'related disciplines' grouped round a problem area? Or is there some sense in which the study of the built environment can arise naturally from the activity of architecture, in such a way as to reconstitute and perpetually renew the intellectual bases on which environmental action and design must be founded? The 1960s opted for the most part for the 'meeting ground' philosophy. We believe the 1970s are turning towards the second answer, and looking into the nature of architecture itself for the disciplines and theory on which the academic study of the built environment can be based. The aim of this paper is to sketch this latter view of architecture as a discipline, not to undervalue the growing interest and involvement of other disciplines in the study of the built environment, but to acknowledge that these contributions will depend for their effectiveness on the evolution of a body of theory and research at the heart of the subject, that is in the nature of architecture itself and the society that produces it.

The 'meeting ground' philosophy is usually supported by a number of apparently powerful arguments. It is pointed out that any breakdown of the study of the environment is inevitably interdisciplinary (as though current disciplinary demarcations and 'cognitive styles' in the academic world were pre ordained); a distinction is often drawn between those disciplines whose outcome is intended to be 'knowledge' (sciences) and those whose outcome is 'action' (professions); and as a consequence, a clear distinction is made between 'science' and 'design' which strongly preserves the identity of each. This convincing paradigm has penalties. In particular, the design or action disciplines tend to remain intellectually weak while the existence of increasing sources of supposed 'knowledge' relevant to design makes their activity more difficult to accomplish.[8] The academic study of design and planning, as well as its practice, becomes subject to every wind that blows in the academic world. Compounded with its subservience to economics, politics and technology, it is not surprising that environmental design does not appear to be guided by a powerful and humane theory as an optimistic society appears to expect.

The disciplinary complexity of environmental studies certainly constitutes a major problem in the development of any uniquely 'environmental' or 'architectural' theory in the sense discussed. It is relatively easy to construct long lists of established disciplines and subject areas which appear to contribute in some way or another to environmental and architectural concerns. Attempts to make sense of the diversity of the subject matter run through architectural discourse from Vitruvius's first chapter on the education of the architect[9] through to some of the most recent 'systems-based' descriptions of the subject area. In fact, it is difficult to avoid the conclusion that environmental studies, at whatever scale of concern, is a 'science of everything'.

Sciences of everything

But environmental design is really not quite so complex as this view suggests. The possibility that this extraordinary total complexity is to some extent a product of our particular way of seeing it – or in other words an artefact – should at least be considered. This may be investigated by looking for analogies.

It turns out that sciences of everything – which are sciences that by this analysis turn out to be connected with everything else in the world – are not so uncommon. A hypothetical example illustrates the point. Imagine a society which had every science we have today except economics, but which had become aware of the need for such a science. How would this society go about constructing the science of economics? It would, of course, construct it out of the disciplines it already had, both for empire-building reasons and intellectual reasons. Sociology, for example, would see the new subject as an aspect of itself, probably as a sub-specialism. Psychology would argue its fundamental contribution because all economic behaviour emanated from the motivation of the individual. Anthropology would argue its foundational role in comparing societies from the point of view of systems of exchange . . . and so on.

All would be justified in their claims. All these disciplines are to some extent represented in this way in the modern science of economics. But none of these contributions would make sense without the existence of economic theory and method to make up the core of the whole subject area. It is probable that our imaginary society would spend much time trying to construct economic theory out of extensions of other disciplines but then would realise that economic phenomena existed in their own right and that it was possible to have a theory about economic phenomena which was relatively independent of other disciplines – perhaps using basic ideas from these disciplines in a rather simplistic way as assumptions to form part of tentative theories, as might be said of economics today. Economics is more or less co-extensive with sociology in terms of the field it covers, but it has a particular way of looking at the phenomena of society. By viewing them in this special way it turns out that the phenomena can be treated as a systematically connected set possessing systemic attributes which may be represented in theory and convincingly related to real-life events. But it is of course economic theory which makes economics independent. This defines its selection of relevant phenomena and a way of interpreting them in a field which may otherwise appear undifferentiated as a general quarry for all-comers. Economics is a 'route through' social phenomena.

Such a situation actually did prevail in the nineteenth century in the scientific subject now called linguistics. According to Trnka, the subject was 'psychologized and atomized' by being regarded 'as a conglomerate of psychology, physiology, sociology and other disciplines'.[10] Structural linguistics succeeded in establishing the study of language as a theoretically independent discipline to such good effect that it constitutes today probably the most powerful body of theory in the human sciences and is currently powerfully colonising the disciplines of which it was previously thought to be an offshoot. All these disciplines are still part of linguistics and we call them psycholinguistics, sociolinguistics and so on, but *not* linguistic psychology or linguistic sociology. (Compare this with the recently established environmental psychology as opposed to a possible psycho-environmentalism.) The difference is subtle but important. The focus is linguistics; these are aspects of linguistic theory and not vice versa.

The concept of an environmental discipline

The linguistic analogy may be used in other ways to discuss the possibility of an environmental discipline. The extraordinary and commonplace thing about language is that everyone can speak. Everyone can produce highly complex series of sounds within complex rule structures to convey complex meaning without *explicitly* knowing the rule structure which governs language. In fact, no one understands the rule structure. The essential structure of a language as a form of rule-using behaviour is still the major mystery in linguistic theory. Nevertheless, we all speak without making many mistakes. (This is, of course, not only a relatively simple question of grammar, but a question of what grammars are and how they may be scientifically described.) The difference between being able to speak and being a linguistic scientist or theoretician is a useful analogy. It is parallel to the difference between learning to use a language-like rule structure in design and making the man-made environment the object of science and theory.

This may be made more explicit. In language, we know that a complex, generative and open-ended rule structure must exist between the domain of structured sounds (phonetic production) and the domain of structured meanings, or to put it another way, between the physical and abstract aspects of language. Using the term crudely, this structure may be thought of as a kind of code, comparable to a code which turns speech into electronic impulses and back again in an information channel. Those who have been trained as designers will be using just such a code (although it was probably never taught explicitly, it was learnt by just being in a school of architecture) which enables the designer to effect a translation from individual, organisational and social needs to physical artefacts. This code which has been learned is supposed to express and contain actual connections which exist between human needs and their artificial environment. In effect, the designer learns to 'speak' a language – to make a useful transition between domains which are unlike each other (sounds and meanings in language, artefacts and needs in design) by means of a code or system of codes which structure that connection. Just as a man who can speak can realise the various functions of language necessary to existence – communication, thinking, ordering, classifying and so on – because he can relate sound to meaning, so designers can realise the functions that society requires of building – climate modification, symbolic expression, resource modification, activity containment – because they can relate needs to artefacts through the code.

A designer uses a code that expresses the connectivity between needs and artefacts in order to make useful and viable links between the two domains. An architectural theorist is concerned to *study that connectivity as it really is and as it expresses itself in the designer's coding structure and in other coding structures used in the interpretation and use of the environment*. It is important to internalise this distinction. It is the difference between being a speaker and being a student of language. It expresses why the preoccupations of the architectural theorist are co-extensive with, and in a sense the same as those of a designer, and at the same time quite different. Instead of using the code to achieve real objectives in the world, we have stepped back to study the code as a social phenomenon and its relation to the real world as it is.

It is on this foundation that a science of environment can be conceived which is based on what designers actually do rather than on the structure of the 'related disciplines'. But this requires a further important distinction, which once again can be introduced by analogy with linguistics. This is the difference between studying a language as a whole and studying the use that individuals make of language in speaking. The environmental equivalent is the morphological history of artificial space and social process on the one hand, and the particular appropriation of it by individuals, groups, organisations and cultures on the other. The environment as a morphological set exists over and above any individual use of it. It exists at any time as a historical 'given'. Those who experience it had no hand in its making, but they will pass it on to the next generation as a 'given' in modified form. These modifications will be the result of particular appropriations; the modified environment will express both these and the previously given structure. Thus, although the overall morphology of the environment and the individual appropriation of it must be considered to some extent separately, they are also interconnected. One of the objects here is to find out how.

Let us summarise this and try to find some useful generalisations. It is usual to represent man and environment as an *interface* touching at all points, but some being more significant than others. It is suggested here that although the interface concept is valuable it belongs properly to the interface between the *artefact* and its environment. The relation of man and environment is not an interface at all but an elaborate structure of relations which has the nature of a code.[11] In this view, the two domains of human needs and physical artefacts have been 'pulled apart' to reveal the structure of connections between them. This is the dominant system of interest for environmental and architectural theory. Designers must use code structures in order to design buildings; people use them to experience it; society constrains designers and users through them. This offers three basic generalisations regarding the study of artificial language-like systems. First, the problem of effecting a relationship between dissimilar domains of entities is very basic to our ways of thinking and even to our mode of existence. We depend on this ability. Second, these operations have a formal resemblance to the idea of 'mapping' between domains in mathematics, a concept which is fundamental to the practical application of mathematics. Third, what is of chief interest to environmental and architectural science are certain kinds of relationships which are in some sense mediated or modified by environmental change. This leads us to the concept of function, seen not in the traditional sense of architectural discourse, but in the scientific sense.

The concept of function

The concept of function is important because it expresses exactly such relationships. In a recent paper, 'The idea of architectural research'[12], this concept was applied to architecture and shown to underlie the organisation and development of architectural research. It was suggested that a useful conceptualisation of a building was the 'four-function model' which identified the building as behaviour-modifier, as climate-modifier, as symbolic-modifier and as resource-modifier. Each of these functions can be expressed as a relationship. The first is between behaviour and spatial structure which is mediated by building; the second is a relation between human psychophysiology and the natural environment, mediated by building; the third is a relation between the physical artefact as a sign and its symbolic meaning, mediated by building; and the fourth is a relation between the use of resources and goals, mediated by building. These relations are in effect realisations of a more basic set of relations. The matrix (Figure 5.1) gives a straightforward version of the structure of these more basic relationships. The columns represent the basic categories – the relation between *man and nature*, and the mediation between *man and man*, which are ubiquitous in all forms of the artificial environment; and the rows represent how buildings function both in a visible, tangible way (buildings as things) and in a less obvious but pervasive way as a cultural language (buildings as signs). The boxes of this matrix yield the four-function model.

In functioning as a climate modifier, the building acts as a complex environmental filter between inside and outside, modifying (by decreasing, increasing, selecting and specifying) the sensory inputs into the human occupants, and also having displacement effects on the external climate. As an activity modifier, the building inhibits some activities and facilitates others, perhaps prompting or determining them. It also locates activities within a broader ecological framework and constitutes a modifier of the total behaviour of that part of society that comes into contact with it.

As a symbolic modifier, the building functions not only in terms of the designer's intentions but also in terms of the expectations and awareness of those who experience it. In this way, the building has a similar displacement effect on societies' symbolic systems as a whole. As a resource modifier the building

	MAN-NATURE relation	MAN-MAN relation
buildings as THINGS	climate modification function	activity-space function
buildings as SIGNS	economic function	social language function

Figure 5.1 Four-function model: *simple matrix version*.

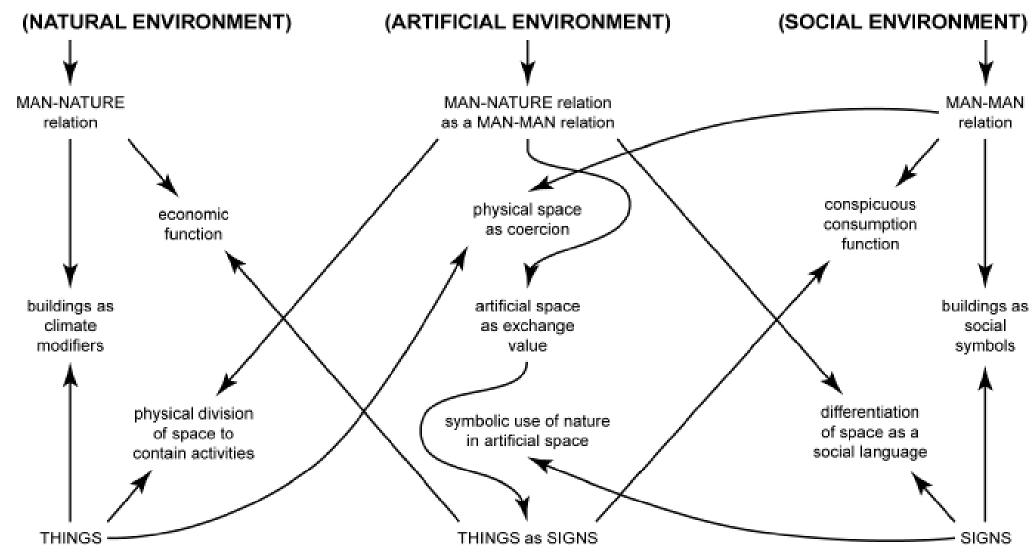


Figure 5.2 Four-function model: *full matrix version*.

functions in transforming existing patterns of use value and exchange value. It adds to the value of its raw materials, acts as a capital investment, redistributes scarce resources of manpower and material over space and time and, in a less measurable way, adds to the existing use value of the building stock.

Figure 5.2 gives a more complex and true-to-life version of these sets of relationships, and shows how, in fact, they embody contradictory pulls. The basic components of the matrix are, as before, man-nature relations, man-man relations, buildings as things and buildings as signs. But further complexity is added by thinking of the *intersection* of similar components in the matrix. This gives the rather unfamiliar idea of man-nature relationships acting as man-man relations (this means, an artificially made, socially constructed surrogate 'nature') and buildings as signs (which means as economic or resource signifiers in themselves). The intersection of like categories in the matrix along with the *union* of unlike categories provides a more complete map of the way in which social space is constructed and used.

Take, for example, the intersection of buildings as things and buildings as signs (things as signs) and its union with nature (man–nature) and society (man–man), respectively. This provides the two major, and contradictory, dimensions of how buildings function economically. The first gives an economic function based on minimum cost in the exploitation of natural resources; the second gives an economic function based on exploitation of prestige which promotes conspicuous consumption. Similarly, if the intersection between nature and society (man–nature relations as man-man relations) is unified with buildings as things and buildings as signs respectively, two types of spatial differentiation are derived.

One is based on the pure physical division of space to contain activities; the other expresses how the division is used as a social language. The other categories in the matrix are derived in a similar way.

As argued in 'The idea of architectural research', the four-function concept has, implicitly at least, underlain the development of architectural research over the past decade. This represents a significant shift from previous thinking about architectural research in which it was presumed that research was the prerogative of the 'knowledge' based disciplines and nothing to do with the 'action' disciplines like architecture and planning. With this model, it is easy to see why no amount of sociological or psychological or engineering or materials research on its own would tell us what, as designers, we really needed to know. The knowledge we require is more about the structure of connections between human needs and physical artefacts as they exist in the real world.

It is through the development of research into such relationships that the possibility of a scientific and integrative approach to architectural and environmental problems, based on the sphere of interest of designers and problem-solvers rather than the academic structure we have inherited, has become possible. It has also had the useful effect of shifting the focus of research from the methodology of design to the nature of the artefact itself, while at the same time making the connectivity of science and design much less of a problem. Environmental design and research can, it appears, converge considerably as far as their subject matter is concerned while retaining the basic distinctions in how they approach it and how they apply it.

Architecture as a science of the artificial

If this approach is useful, then what are the implications for the old conflict between architecture as art and architecture as science? Our argument is that there is no conflict, any more than there is a conflict between using and studying language. No linguistic research has ever had a bad effect on language. The fallacies of 'basic English' arose not from linguistic theory but from linguistic scientism. Linguistic theory today takes off from the concept of a 'rule-governed creativity' (to use Chomsky's phrase). In fact, its theories today are essentially answers to the question: how can unconscious organised systems permitting 'rule-governed creativity' exist in our heads, what are they like and how do they get there? This is exactly the foundational concept that we need in architecture.

Using this basic idea, and orientating research towards an understanding of unconscious social codes that construct both our awareness of space and our actions as designers, architecture can become a member of the community of truly modern sciences without sacrificing anything of its preoccupation with the human, the intuitive, and the free run of the socio-spatial imagination. But there is a problem. If architecture sustains such a theoretical enquiry into its own foundations, in society and in the evolutionary morphology of built form, then it will no longer be the same type of system as it was before. Its consciousness of itself, its increasing ability to externalise its social nature and effect, will itself be an important component in the system under study. We arrive at the problem of sciences that are part of their own subject matter.

Again, this is not unusual. Returning to economics, a society which has an economic theory, and maybe mathematical models of the economy based on that theory, is not the same society that it was prior to having that theory and those models. The model may itself become the most important single aspect of the economy. Today, the economy fluctuates as a result of its own contemplation of itself, made possible by economic science. In the field of human endeavour, scientific theories become part of the society in which they develop. This is not new. It is this that gives social reality its abstract nature. It has been true in economics ever since 'Aristotle discovered the economy',[13] and theoretical ideas began to construct the role of the market in urban societies.

The proper name for these sciences which become part of themselves and create their own universe should not be 'human' or 'social' sciences but 'sciences of the artificial'. Their essential concern is with the self-developing, self-perpetuating properties of man-made systems like languages, cities, economies and even society itself. The 'social and human' sciences concept is insufficient because it is implicitly constructed from a direct analogy with natural sciences. The concept of 'sciences of the artificial' arose through increasing awareness that within a decade of their development, computers had to become the subject of a 'natural history' in order to understand what was happening to them and how they were changing the environments in which they were out.[14] Architecture is a science of the artificial, and its mode of action-design – as well as its unconscious codes and theories – are part of its subject matter, and not some rootless extension into an ethereal domain of unconstructed intentions and ideals.[15]

About the authors

At the time of publication of this paper in 1976, Bill Hillier was Head of the Research Section of the Royal Institute of British Architects and Senior Research Fellow at the Unit for Architectural Studies, University College London. Adrian Leaman was Research Officer at the RIBA.

Notes

1. B. Hillier and A. Leaman, 'A new approach to architectural research', *RIBA Journal – Royal Institute of British Architects* 79(12) (1972): 517–521; B. Hillier and A. Leaman, 'How is design possible?', *Journal of Architectural Research* 3 (1974): 4–11; B. Hillier and A. Leaman, 'Architecture as a discipline', *Journal of Architectural Research* 5 (1976): 28–32; B. Hillier, A. Leaman, P. Stansall and M. Bedford, 'Space syntax', *Environment and Planning B: Planning and Design* 3 (1976): 147–185.

2. B. Hillier and J. Hanson, *The social logic of space* (Cambridge: Cambridge University Press, 1984).
3. Vitruvius. *The ten books on architecture*, translated by M. H. Morgan, with illustrations and original designs prepared under the direction of H. L. Warren (New York: Dover, 1960).
4. J. Jacobs, *The death and life of great American cities* (Harmondsworth, Middlesex: Penguin, 1961).
5. P. Panerai, J. Castex and J. C. Depaule, *Formes urbaines, de l'îlot à la barre* (London: Architectural Press, 1975).
6. S. Anderson (ed.), *On streets* (Cambridge, MA: MIT Press, 1978).
7. C. N. F. Santos, *Movimentos urbanos no Rio de Janeiro* (Rio de Janeiro: Zahar, 1981).
8. B. Hillier, J. Musgrove and P. O'Sullivan, 'Knowledge and design', in *AIA-Architect-Researchers' Conference. Environmental design: research and practice. Proceedings of the Third Conference of the Environmental Design Research Association*, edited by W. Mitchell (Los Angeles: University of California Press, 1972), pp. 29-3-1 to 29-3-14.
9. M. V. Pollio, *The architecture of Marcus Vitruvius Pollio, in ten books*, vol. 128 (London: Lockwood, 1874).
10. B. Trnka, 'Prague structural linguistics', in *Structuralism: A reader*, edited and introduced by Michael Lane (London: Cape, 1970), pp. 73–84.
11. B. Hillier and A. Leaman, 'The man–environment paradigm and its paradoxes', *Architectural Design* 43 (1973): 507–511.
12. Published as Hillier and Leaman, 'A new approach to architectural research'.
13. K. Polanyi, 'Aristotle discovers the economy', in *Primitive, archaic and modern economies: Essays of Karl Polanyi*, edited by G. Dalton (London: Anchor Books, 1968), pp. 78–115.
14. H. Simon, *The sciences of the artificial* (Cambridge, MA, and London, England: MIT Press, 1969).
15. This paper was originally prepared in April 1974 and presented to a conference at the York Institute for Advanced Architectural Studies concerned with activity and space. The authors would now give equal emphasis to the study of built form itself as to function and their current research is concerned with the development of a theoretical model for the description of space relationships in buildings. B. Hillier, and A. Leaman, 'The architecture of architecture: Foundations of a mathematical theory of artificial space', in *Models and systems in architecture and building*, edited by D. Hawkes (Hornby: Construction Press, 1975), pp. 5–28.

6 Space syntax (1976)

Introduction to 'Space syntax'

Sam Griffiths

The title of the 1976 article 'Space syntax' by Bill Hillier, Adrian Leaman, Paul Stansall and Mick Bedford is unusual, for an academic publication, in its brevity. A subtitle would have given the reader a greater steer on its contents. We may speculate that there was a reason they chose not to add one. Bill Hillier had been immersed in the philosophy of language since studying English Literature at Cambridge and understood the sciences of the poetic as much as architectural artifice. He would have been well aware that to deploy just two short alliterative words comprising three equally stressed syllables, 'Space syntax', in the title was to add dramatic emphasis. For Hillier and colleagues this was not just another article but rather an *announcement* of something new. Any qualifying subtitle would have diluted its impact, hence 'Space syntax' stands alone.

Space syntax was Hillier and colleagues' response to the problem of the *description* of complex built forms seen as the collective products of human endeavour. In the absence of a satisfactory independent description, they argued, human settlements could never be recognised as socially generative in themselves but only as representations projected by socio-economic or cultural social structures. The authors were strongly influenced in their 'theory of patterns' by Newton's rejection of the Aristotelian *telos* that sought to assign an essential socio-economic or cultural purpose to built forms.

The article elaborates the syntax of the 'morphic language' of human spatial organisation. This is defined first by combinatorial rules governing a given set of objects, relations and operations. Second, through their restriction on a random process of planar growth. Third on the aggregation of elementary objects (the 'minimal setup') within a given carrier space on the surface of the earth. The authors note that the 'problem of describing structures, especially collective structures . . . exists over and above the problem of generating them'. This is because the global forms settlements take are not wholly knowable on the basis of the local rules that generate them. The article consequently focuses on the generative rules that are hypothesised to link the global settlement descriptions we see within the cartographic record to the social-historical process of their construction.

Readers who are not specialists in linguistics or who fear they may struggle to digest the esoteric notation (or 'ideography') of space syntax should be reassured that it is a surprisingly *intuitive* proposition. Unlike the rules of mathematical logic but similar to those governing meaningful verbal expression, the morphic language is said to be constrained by 'being realised in the experiential world [and] being creatively used for social purposes'. Unlike natural languages (but similar to mathematical ones), the morphic language does not signify anything beyond itself; it 'constitutes patterns which are their own meaning'. This is why space can be truly said to be *social space*; it is because humans find the inhabited world *intelligible*, at least in some minimal sense. On the basis of a wide survey of settlement types, the authors make the remarkable claim to have established eight generative syntaxes that account for the spatial morphologies of recorded settlement forms. They distinguish between two major types: 'distributed syntaxes' that describe systems characterised by the arrangements of individual elements in open space – each space playing an equal part in building the global space pattern; and 'nondistributed syntaxes' that describe systems characterised by boundaries arranged hierarchically, with some spaces playing a controlling part in building the global pattern.

Publishing 'Space syntax' in *Environment and Planning B*, then the recently launched journal of architectural and buildings research edited by Lionel March, was an important milestone for the collaboration of Hillier, Leaman, Stansall and Bedford. After years spent honing their ideas in niche journals it brought their work to a wider interdisciplinary audience. The article was later reprinted in its entirety in 1978 as a contribution to the British Archaeological Reports series. This publication features a spirited correspondence between the authors and the renowned British anthropologist Sir Edmund Leach pivoting on whether verbal language was, as Leach argued, the most appropriate intellectual framework for understanding the sociality of historical settlement forms. The advocates of space syntax as a *morphic* language, of course, saw things differently.[1]

The authors' position on morphic languages would eventually receive clarification in Bill Hillier and Julienne Hanson's definitive 1984 book *The social logic of space*.[2] The opening chapter of this work, 'The problem of space', offers a refinement of the propositions advanced in 'Space syntax'; itself representing the maturation of an intellectual agenda first sketched by Hillier and Leaman in their 1974 paper 'The architecture of architecture'.[3] The shared authorship of the 1976 article serves as an invaluable reminder that posterity should not abbreviate the intellectual history of a field to the individual with which it is most associated, however influential Bill Hillier became in driving the development of space syntax research over subsequent decades. It is important that not only Adrian Leaman, who co-authored a number of key early papers with Hillier, but also Paul Stansall and Mick Bedford should be acknowledged as pioneers of the field from its earliest days – architects and thinkers whose contributions were formative.

For a scholar working in the field of contemporary space syntax research, re-reading of 'Space syntax' is like returning to the intellectual source for nourishment. It prompts reflection on how widespread interest in the potential of space syntax 'tools' of spatial analysis and data visualisation should not (and need not) come at the cost of neglecting its core theoretical proposition: that the architectural spaces of human life are fundamentally *social* in their material descriptions and *socialising* with regard to their implications for human action. It can also help stimulate new research agendas and revive older ones. Generative syntaxes have clear contemporary relevance in the use of computational techniques to explore novel possibilities in architectural forms on the basis of parameters suggested by the historical and archaeological record. The theory of morphic languages deserves further research and development as an as yet relatively unexplored branch of social ontology; and there is still much empirical ethnographic and historical work to do in understanding the relationship between 'spatial and social syntax'.

Space syntax

Bill Hillier, Adrian Leaman, Paul Stansall and Michael Bedford

Originally published: Hillier, B., Leaman, A., Stansall, P. and Bedford, M. 1976. 'Space syntax'. *Environment and Planning B: Planning and Design*, 3, 147–185.

This paper addresses itself to the question of how and why different societies produce different spatial orders through building forms and settlement patterns. It consists of three parts. Firstly, at a meta-theoretical level, it is suggested that spatial organisation should be seen as a member of a family of 'morphic languages' which are unlike both natural and mathematical languages but which borrow properties from each. In general, morphic languages are used to constitute rather than represent the social through their syntax (that is the systematic production of pattern). Secondly, a general syntactic theory of space organisation is proposed. It is argued that spatial patterns in both complex buildings and settlements fall into eight major types, which are interrelated in structural ways. Finally, the syntactic theory is used to integrate a number of recent general propositions made in anthropology regarding human space organisation.

1 Preliminaries: Mathematics and the 'sciences of the artificial'

> 'We cannot understand the flux which constitutes our human experience unless we realise that it is raised above the futility of infinitude by various successive types of modes of emphasis which generate the active energy of a finite assemblage. The superstitious awe of infinitude has been the bane of philosophy. The infinite has no properties. All value is the gift of finitude which is the necessary condition for activity. Also, activity means the origination of patterns of assemblage, and mathematics is the study of pattern.'[4]

Belief in a mathematical order inherent in nature has always been a fundamental postulate of theoretical science. First put forward by the School of Pythagoras, which developed a numerical theory of natural order from such discoveries as the relation between musical harmonies and numerical proportions, it was linked by Galileo to the experimental method, and together they form the dual foundation of the modern conception of science. Analytical geometry, calculus, group theory, non-Euclidean geometries and perhaps catastrophe theory were all subsequent steps in linking our conceptions of natural order with mathematics. However unreasonable a belief mathematical order in nature may appear in principle, the 'unreasonable effectiveness' of mathematics in the natural sciences leaves no doubt that it has been amply justified by events.

But the sciences of man-made entities like settlement patterns, societies and languages have no such record of success to confound the sceptic. Moreover, the claims of these sciences to be excused for their poor mathematical development on account of their extreme youth sounds more and more uneasy as decades pass. However, the root reason for the lack of mathematical theories in the 'sciences of the artificial' may be that they are not sought after, since the fundamental postulate justifying the intervention of mathematics in these sciences is not a belief in a mathematical order inherent in the objects of study, but simply a belief in the power of mathematics as an instrument. In principle, such a reduced claim appears justified. Even if nature does work mathematically, this does not imply that man the artificer also does. To believe in a mathematical order inherent in complex artificial entities requires us to believe that man creates more mathematically than he knows.

This argument strongly resembles the one that eventually led to Galileo's condemnation. He was required by the church to believe that mathematical models were only convenient instruments for describing and predicting nature, not expressions of an order present in nature herself. One suspects that today a similarly modest instrumentalism would be more acceptable to the high priests of the unnatural sciences than a Galilean belief in inherent mathematical order.

There is, however, one scientific problem area which suggests that the sciences of the artificial may be forced into thoroughgoing Galileanism. All branches of artificial intelligence research have encountered a major conceptual barrier – the problem of representing fields of knowledge. 'Machine intelligence is fast attaining self-definition, and we now have as a touchstone the realisation that the central operations of intelligence are (logical and procedural) transactions on a knowledge-base.'[5] This problem, according to Michie, is now the common denominator of research into artificial pattern recognition, machine translation, and even chess playing, of which Michie wrote: 'As with other sectors of machine intelligence, rich rewards await even partial solutions to the representation problem. To capture in a formal descriptive scheme the game's delicate structure; it is here that future progress lies, rather than in nanosecond access times, parallel processing, or mega-mega-bit memories.'[6]

It is hard to see how such problems will be resolved except by novel theories of combinatorial pattern formation. If mathematics is to justify its claim to be the general abstract study of pattern, such theories will be assimilated to mathematics, if they are not already part of it. We cannot know in advance whether the new combinatorial ideas we need will come from mathematics, or whether they will come from outside and challenge mathematics, as physics has done so often. It may, for example, be the case that modern mathematics in pursuit of its most abstract and intangible foundations has neglected certain simpler, perhaps more imperfect, types of order than may prevail locally in ordinary space-time. If this were so, it would at least allow us to adopt a Galilean position in the sciences of the artificial without an extravagant belief in a relation between the most unworldly domains of abstract mathematics and the everyday world of practical pattern recognition, language analysis and so on.

Whatever the solution, the existence of the knowledge problem in artificial intelligence research already strongly suggests that some formal order, of a more or less mathematical kind, must inhere in the complex entities which we 'recognise' so easily in everyday life. With only a minor extension of the argument, it may be suggested that this could be a key to the scientific study of those artificial systems which are defined on a collectivity, like cities, societies and languages, and depend on continuous, largely unconscious pattern recognition by members of that collectivity. In this perspective a key relationship comes to the fore: the relationship between the formal structure of what there is to be known (for example, the patterns of space organisation, patterns of social networks and so on); and the formal mental structures by which these are known or recognised. It is then an obvious hypothesis that the same formal structure could account for both.

This extension of current debates in artificial intelligence also suggests an alternative strategy for research to be conducted in parallel to the problem of formal representation of knowledge fields on the computer: that of the analysis of artificial systems like space patterns and social patterns for inherent formal structures which might contribute to their knowability.

These are the theoretical considerations that underlie the 'space syntax' research programme. They are intended both to apologise for adopting a formalistic approach to phenomena not commonly thought to be responsive to such an analysis, and also to apologise for the adoption of a practical and empirical rather than purely mathematical approach. Our aim has been to work towards mathematisation from intuitive formal principles, rather than to adopt a branch of mathematics such as topology or graph theory and work towards the phenomena. Experiments have been made in moving more firmly in the direction of one or other branch of mathematics, but each time it has become clear that this would impose severe limitations on our ability to stay close to the real evidence, and to try to draw its formal structure from it.

Perhaps a firmer argument for a 'syntactic' rather than properly mathematical strategy is that, even within the scope of a general belief in an inherent formal order giving rise to knowability in space patterns, we cannot know in advance which of the array of current branches of mathematics will be appropriate, or even if any branch will offer models for the level and type of approximation we require. The proper scientific strategy therefore seems to be to build a theory of patterns, with a close respect for the evidence but without too much regard for early justification in mathematical terms. Although we realise that we are bound to be strongly criticised for our neglect of mathematics, we hope we may be excused on the grounds that our resulting model of the formal syntax of human space organisation is at least 'unreasonably effective' in characterising the space patterns made by human societies, in showing how they were generated, how they relate to social patterns, and perhaps above all, in showing how even the most complex patterns are 'knowable' through knowledge of a few elementary concepts and operations.

In brief, our hope is to have made an effective model of the 'knowledge-field' constituted by architectural and urban space patterns. But we have done so at the price of mathematical acceptability. We therefore ask the reader to consider three things. First, the match between the model and the empirical evidence; second, the internal consistency of the model within its own limited syntactic terms, rather than in terms of its agreement with basic mathematical ideas; and third, the possibility that man-made systems involving patterns of relations, especially those defined as collectivities such as human populations, or collections of spatial domains, may *require* this syntactic level of formal analysis to mirror their real internal patterns, rather than the more searching analysis of mathematics proper. It may even be ventured that progress in developing formal theories of complex artificial systems is handicapped by not having such a level of formal analysis.

In view of the centrality of a 'syntactic formalism' to our whole research enterprise, we have developed it into an explicit theory, the theory of *morphic languages*, which appears to us to fall between mathematics and natural languages, and to offer the appropriate general concept for the analysis of complex artificial systems involving patterns defined on collectivities. As such the theory of morphic languages, and the worked example given – the morphic language of space patterns – are offered also as a contribution to the study of 'collective phenomena'.

2 The theory of morphic languages

If the problem of *knowability* is defined as that of understanding how characteristic patterns in a set of phenomena can be recognised by reference to abstract principles of arrangement or relationship, and the problem of *morphology* is defined as that of understanding the objective similarities and differences that a set of phenomena characteristically exhibit to ordinary experience, then the aim of the theory of morphic languages is, for certain classes of real, socially defined collective phenomena like spatial patterns, to unite the two problems into the single problem of understanding how the morphology may be generated from a parsimonious set of elementary objects, relations and operations. In effect, the reduction of morphology to the elementary structure of a combinatorial system is argued to be its reduction to its principles of knowability. The set of combinatorial principles we call syntax. Syntax is the most important property of a morphic language. What is knowable about the morphological output of a morphic language is its syntax. Conversely, syntax permits the morphology to exhibit regularity in its similarities and differences.

Syntax in a morphic language is defined as a set of related rule structures formed out of elementary combinations of the elementary objects, relations and operations. These can be introduced, independently or conjointly, in a *minimum setup* for the morphic language to produce recognisable patterns. A minimum setup is a morphic language without its syntax, that is to say, a morphic language operating randomly. More exactly, a minimum setup consists of a space[7] within which the morphic language can operate (that is, generate patterns) called the carrier space; a minimal rule of operation, that of simple repetition at random intervals; a minimal object, the least complex permitted by the system; with minimal relations, where each object has only the relation of belonging to the carrier space. It is reasonable to call such a setup 'random' since each event (that is the placing of an object) that takes place in the carrier space is independent of every other event except in that all belong to the same setup. This follows the definition of a 'chance setup' by Hacking.[8]

A morphic language therefore consists of:

a. *a minimum setup*, made up of a carrier space and a randomised ongoing process;
b. *a syntax*, that is a set of elementary objects, relations and operations capable of being combined to form rule structures to restrict the randomness of the minimum setup; and
c. *a syntax-rule* – a rule for the formation of rules – which ideally should exhaust itself against some natural or logical limit. For example, in space syntax, the rule exhausts itself against the barrier of not being able to develop beyond three-dimensional space (in fact slightly earlier: against the fact that people cannot fly. Man-made space is effectively two-dimensional because movement is two-dimensional. Stairs are a two-dimensional reduction of a three-dimensional reality).

A morphic language links the foundation of probability (the 'chance setup') with the fundamental idea of a mathematical structure (objects, relations and operations) from the outset. The advantages of the morphic language concept are several. First, by linking a probabilistic approach with a structural approach from the start in the modelling of phenomena, order and pattern are seen as improbable, being the result of the introduction of syntax into the minimum setup. Second, it offers, in principle, a method of keeping a record of all the order that has been built into a system. A morphic language is not decomposable into 'sub-systems', but only into the syntactic rules responsible for the production of this or that kind of pattern. Even with all syntactic rules removed, the minimum setup is still, as will be shown, a relatively rich system, in spite of being minimally ordered. Third, randomisation plays a key part in the formation of certain major pattern types that appear in the real world of space patterns. Certain patterns are only produced by a generative process if the process is randomised apart from its syntactic rule (for an example, see the global pattern that emerges in generating a 3-syntactic surface in Figure 6.9). Fourth, a morphic language deals with the many minimally ordered situations that exist, in terms of a theory of patterns, that is, we may treat randomness as a special case of pattern. This turns out to be critical to the problem of relating spatial to social patterns (see section 4 – Space and society). Fifth, a morphic language gives natural, rather than arbitrary, limits: at the lower limit is randomness, at the upper limit are the limits of realising new combinations in real space. In essence, space syntax is an exploration of the combinability of open and closed discs, and open and closed rings (assuming these to be projections onto the plane of open and closed balls, and open and closed tori). Sixth, a morphic language makes the model self-contained in that it can be shown, in the case of space syntax at least, that all the objects, relations and operations that form the syntax, and even the syntax rule itself, can be found by an analysis of the minimum setup.

Intuitively, the postulation of a minimum setup as the basis for a morphic language means that it is assumed that people and societies deploy themselves in space, and that these deployments are capable, under certain conditions, of adopting certain patterns. The research task is, therefore, *not* to say why people deploy themselves in space, but to offer a theory of the patterns. Conceptually this is comparable to the introduction of the inertia postulate into physics. It liberates us from the Aristotelean 'essences' of universal behavioural principles, which plague current theorising about space, and allows us to build a theory of the characteristic space patterns that different types of society and organisation create.

In the theory of morphic languages, therefore, the Newtonian postulate of an ongoing system is added to the Galilean hypothesis of a formal structure inherent in the patterns of order exhibited by the states of the system. However, this formal order is not in the proper sense a mathematical order, but a *syntactic* order. But the word syntax is not used with the same technical specification or theoretical status usually assigned to it in linguistics. It is therefore necessary, to avoid misunderstanding, to be clear about the relation between a morphic language and mathematics on the one hand, and natural language on the other.

The primary purpose of a natural language (irrespective of particular linguistic 'functions') is to *represent* the world as it *appears*, that is to convey a *meaning* which in no way resembles the language itself. To accomplish the task of representation in an infinitely rich universe, a natural language possesses two defining characteristics. First, a set of *primary morphic units* which are strongly *individuated*, that is each *word* is *different* from all other words and *represents different things*; and second, a *formal* or *syntactic structure* which is *parsimonious* and *permissive*, in that it permits infinitely many sentences to be syntactically well-formed which are semantically nonsense (that is, effectively nonsense from the point of view of linguistic form as a whole). Conversely, meaning can be transmitted (that is represented) without well-formed syntactic structure in certain cases.

The defining characteristics of a natural language are a relatively short, possibly conventional, grammar, and a large lexicon.

By contrast, mathematical languages have very *small lexicons* (as small as possible) and very *large 'syntaxes'* in the sense of all the structure that may be elaborated from the initial, minimal lexicon. Such languages are virtually useless for representing the world as it appears because the primary morphic units are not individuated at all, but rendered as *homogeneous* as possible – the members of a set, units of measurement and so on. Mathematical symbols strip the morphic unit of all its particular properties – of being a member of a set, of existing and so on. To be interested in the particular properties of particular numbers is, for a mathematician, the equivalent of a voyage into mysticism. Mathematical languages do not *represent* or mean anything except their *own structure*. If they are useful for representing the most abstract forms of order in the real world, it is because, in its preoccupation with its own structure, mathematics arrives at *general principles of structure*, which, because they are deep and general, hold also at some level in the real world.[9]

Morphic languages differ from both, yet borrow certain properties from each. From mathematical languages, morphic languages take the *small lexicon* (that is the homogeneity of its primary morphic units), the *primacy of syntactic structure* over semantic representation, the property of being built up from a *minimal initial system* and the property of *not meaning anything except its own structure* (that is to say, it does not exist to represent other things, but to constitute patterns which are their own meaning). From natural languages, morphic languages take the property of being *realised in the experiential world*, of being *creatively used for social purposes* (or permitting a 'rule-governed creativity') and of being *constitutive, rather than representative, of the social*.

Thus in a morphic language, syntax has a far more important role than in natural language. In natural language, the existence of a syntactically well-formed sentence permits a meaning to exist, but neither specifies it nor guarantees it. In a morphic language, the existence of a syntactically well-formed pattern itself guarantees and indeed specifies a meaning, because the 'meaning' is only the abstract structure of the pattern. Morphic languages are the realisation of abstract structure in the real world. They convey 'meaning', not in the sense of representing something else, but only in the sense of constituting a pattern. Thus if, as we believe, both space organisation and social structures are morphic languages, the construction of a social theory of space organisation becomes a question of understanding the relations between the principles of pattern generation in both.

This does not mean that architectural and urban forms are not used to represent particular meanings, but it does argue that such representation is secondary. To achieve representation of 'meaning' the morphic language of space does so by behaving as a natural language. It individuates its morphic units, making them as different as possible from other morphic units. Hence buildings which are intended to convey particular 'meanings' do so by the addition of idiosyncratic elaboration and detail – decoration, bell-towers and so on. In so doing, the morphic units come to behave more like particular words in natural language. Conversely, when natural language is useful to convey abstract structure – as, for example, in academic monographs – it does so by increasing the importance of syntax over the word (see Bernstein's concept of an elaborated code, and its syntactic effects[10]). This is why scientific interest is usually bought at the price of boredom. We cannot, alas, be poets and write scientific papers in extended metaphors.

Morphic languages are also like mathematics, and unlike natural language, in that they pose the problem of the *description*, in addition to that of the *generation*, of structure. Current linguistic theory would assume that a theoretical description of a sentence would be given by a formula expressing generative and transformational rules. This would hold even if current efforts to build semantically (as opposed to syntactically) based theories were successful. In mathematics, however, structure is only reducible to generation if one takes a strong philosophical line opposing the 'reification' or 'Platonisation' of structure, and arguing that all mathematical structure is self-evidently reducible to an orderly activity by mathematicians, not to be thought of as existing in its own right.

Whatever the solution to this problem in mathematics, in morphic languages it can be clearly shown that cases exist where the problem of describing a structure which exists objectively in the real world is over and above that of understanding how it is generated. To take a simple example, imagine that a series of individuals build square, single cell, single entrance dwellings by each joining his cell facewise onto an existing wall in the collection, keeping the growing collective object as compact as possible (that is, ensuring that the largest square obtainable by projecting the lines of cell walls is as small as possible). Given that each individual follows no other rules, then the result will be as follows (see Figure 6.1):[11]

The aggregate object takes the form of a set of cells grouped around courtyards, in most cases of unit or twice unit size. In other words, although each individual only follows a *local* rule, relating only to his cell and the cell to which he joins facewise, the *global* object has a richer, emergent structure not thought of by any individual. Thus the aggregate object is not satisfactorily described by

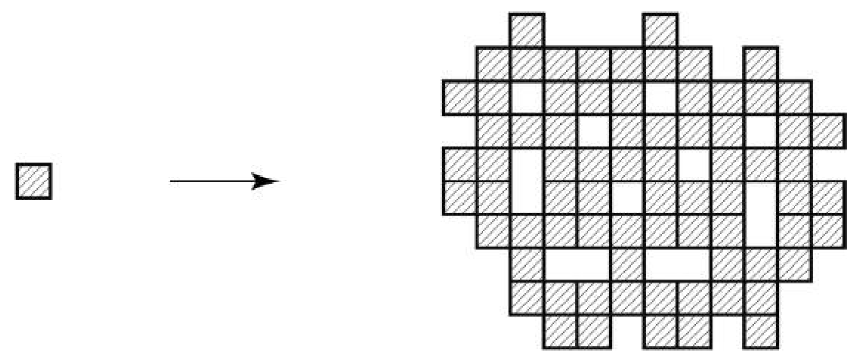


Figure 6.1 Cell – and aggregation of cells.

knowledge of how it was generated. The problem of describing structures, especially collective structures, thus exists over and above the problem of generating them.

In fact, the *dialectic of generation and description* appears to be of fundamental importance in the real-world behaviour of morphic languages. Any ordered collective activity which is not fully preprogrammed gives rise to the problem of *retrieving a description* of the collective pattern. 'Meaning' can be seen as a stably retrievable description. The problem of understanding the growth of cities today is such a problem. Indeed, it might be argued that the role of intellectuals in society is the retrieval of descriptions. An analogy can be made with Arbib's concept of a biological self-reproducing machine[12] which does not contain a permanent description of itself, but which has a facility for retrieving, at any time, a description of its genotype.

In summary, morphic languages:

- are built up from a small elementary lexicon;
- realise syntactic structure in the real world;
- do not 'mean' anything except their own syntactic structure (in which their social purpose resides);
- have both a generative and a descriptive mode.

This paper is concerned with the problem of generation only.

3 The morphic language: 'Space syntax'

In the morphic language 'space syntax' the minimum setup consists of:

a) A carrier space which is the surface of a sphere (the surface of the earth or equally some landmass on the surface of the globe); and
b) An ongoing process of production consisting of:
   - some way of marking sufficiently small parts of the surface so that they are recognisably different from neighbouring parts (if, for example, the surface was uniformly white, it would be sufficient to paint the small parts black; or to place stones would be an alternative);
   - the occasional repetition of such markings; but
   - without relation between one marking and the next (that is, randomly, the only relation being that each mark belongs to the carrier space).

This setup is sufficiently rich to derive by analysis all the objects, relations and operations that constitute 'space syntax'. The 'surface', seen experientially, consists of two kinds of entity: a solid entity – the earth itself; and a vacant entity – the space we can move about in on the solid entity. This corresponds to our general experience of space. The surface consists of continuous or vacant parts, through which

movement is possible, and parts occupied by objects, which impede movement and which we therefore call discontinuities. These are the two elementary *objects* of 'space syntax'. The concept of a continuous space we refer to by the letter 'c', and a discontinuous space by the letter 'd'.

In the minimum setup, a form of demarcation that consisted only of marking the surface so that it is different from its neighbourhood (for example, painting it) would not affect the continuity on the surface, whereas the placing of a stone introduces a local discontinuity. Demarcations can therefore be in a continuous or discontinuous mode. Whichever of these is adopted, however, it is already implied in the minimum setup that each demarcation is finite and independent. Finiteness can be expressed as the relation of being completely contained by a neighbourhood, whether this neighbourhood is itself a continuity or a discontinuity. The relation of containing we write $\supset$, such that whatever is on the left of the sign contains whatever is on the right. Thus the expression $\supset c$ expresses the finiteness of a segment of continuous space, it not being necessary to know what contains the space, only that it is contained. Likewise, $\supset d$ means a finite solid object, or discontinuity.

It has already been observed that it is a natural property of continuous space to be permeable, whereas a discontinuous space is impermeable. The relation of permeability we write $\rightarrow$, meaning that whatever is on the left of the sign has a direct path to whatever is on the right. The expression $\rightarrow c$ therefore means that the space is permeable to whatever is on the left of the sign, and $\rightarrow d$ means that there is impermeability.

The demarcation of a carrier space by means of finite c-objects we call *differentiation*, and that by finite d-objects we call *distinction*. Each can have the property of being *clear* or *unclear*. By this we mean that it is possible to know that a differentiation, $\supset c$, is a finite differentiation without knowing exactly where the limits of the differentiation are. For example, if a white surface is demarcated by clustering black dots, so that there is a zone around the densest black dots where black dots and white areas are mixed, but eventually, sufficiently far from the centre of the cluster, there are no more black dots, then the concept of finite containment is as adequately expressed by $\supset c$ as if the limits were at all points perfectly clear. There is, however, a way of expressing the concept of clearness (as adequately as the simple expression $\supset c$ expresses unclearness) with the concepts and symbols so far established, namely to allow an object to refer the idea of containment to itself. This is already present in natural language in the term 'self-contained' – on reflection a rather odd expression – but meaning that a space-object is within its own limits. The relation of self-containment, which is argued as equivalent to clearness, can be written $\supset c$. The same argument applies naturally to the $\rightarrow$ relation. A space, c, is known to be permeable in its natural state, but it has no clear paths. Paths are, as it were, nonspecific and universal in the space. The expression $\vec{c}$ means that a clear path is introduced into the space – that is, there is some set of marks which differentiate a path clearly from the rest of the space. Likewise, if the self-referring arrow is introduced to the idea of a d-object, $\vec{d}$, then it expresses the idea of a permeable discontinuity, which implies a path that passes through a d-object, not through a space, or c-object. This concept is called the nontraversing path.

The two objects, c and d, and the two relations $\supset$ and $\rightarrow$, permit the construction of an elementary lexicon, consisting of all possible elementary permutations of the terms. This is given in Figure 6.3, below. It should be noted that none of these concepts are viable as independent entities. The purpose of the lexicon is to exhibit the meaning that will be given to these terms when they occur in more complex combinations in syntactic formulae. In the syntactic formulae there is always a $\supset$ sign and a $\rightarrow$ sign to the left of the formula, meaning that there is always a surrounding carrier space and a 'carrier path' (usually implied by the carrier space), since whatever is to be created in the carrier space must be finite and accessible.

So far, two objects and two relations, together with all their elementary permutations, have been derived from an analysis of the minimum setup. Closer inspection, taking into account the elementary lexicon, reveals a more complex phenomenon. There are two forms of continuous object in the setup, those without a hole and those with at least one hole. The object without a hole is the simple c- or d-object as described. The object with at least one hole is its neighbourhood. The insertion of a finite object in the carrier space has the effect of creating a local subspace of the carrier space, possessing a hole, namely the place where the object is. The neighbourhood space, with its characteristic form, is a consequence of the placing of the object. The neighbourhood space is an emergent structure. Moreover, since not the whole of the remainder of the carrier space (recalling the 'sufficiently small' condition for the c- or d-object in the minimum setup) is the neighbourhood, then the neighbourhood must be finite. We need not be discouraged from a belief in its finiteness by our ignorance of where its limits are. This only makes it unclear, and we know already that unclearness does not interfere with finiteness. Since the neighbourhood both has a hole and is finite, then we know that it must approximate the form of a *ring*, or annulus. Thus continuous objects in the form of discs and continuous objects in the form of rings both exist in the minimum setup (see Figure 6.2).

If we consider each as an independent form, then it can be seen that not only is the ring an emergent structure of the disc, but also that the disc is an emergent

Figure 6.2 (a) Disc versus (b) ring.

structure of the ring. Furthermore, if d is substituted for c in the case of the ring (that is, if the ring is a d-object), then the object is a closed boundary, or enclosure (see Figure 6.3).

The enclosure together with its emergent interior disc is a more complex structure than we have so far dealt with. It consists of a discontinuous component and a continuous component in a relation in which the discontinuity contains the continuity. It is a disc with a boundary. This is a fundamental relation, and permits us (at the risk of offending topologists, who will recognise the definition, but not excuse the adaptation made of it for this purpose) to introduce a definition: an object is *closed* if it has its own boundary; otherwise it is *open*. We can therefore distinguish between a closed and an open disc (see Figure 6.4a).

The distinction between open and closed space objects, together with the syntactic rules for constructing open and closed objects, is perhaps the most fundamental in space syntax. The property of being closed is *not* the same as that of being *enclosed*. For example, consider a group of contiguous closed discs grouped around a central space (a 'courtyard' form – see Figure 6.4b). On the given definition, the central space is an *open* disc, in spite of being completely enclosed, since all the boundaries belong to the surrounding discs, not to the central space. In contrast, we have the 'house and surrounding garden' form. The garden is a closed disc (although 'open' with respect to the 'house') since it has its own boundary

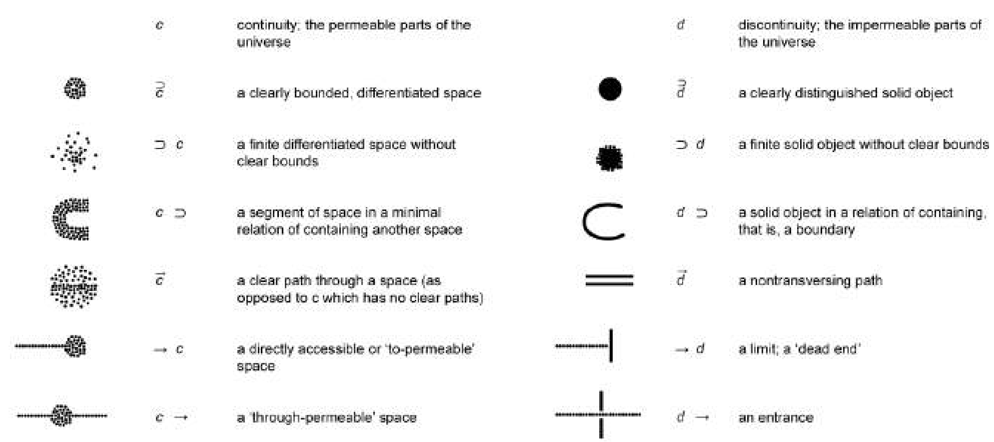


Figure 6.3 Elementary lexicon with illustrative diagrams.

Figure 6.4a A closed versus an open disc.

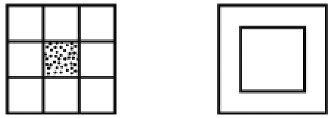


Figure 6.4b A 'courtyard' form versus 'house and surrounding garden' form (left and right).

(see Figure 6.4b). By the same argument, streets are 'open' spaces, whereas the allegedly 'open' spaces of the average 'estate' are in fact closed.

The purpose of this preliminary analysis of the minimum setup has been, first, to show that certain basic forms are syntactically inevitable once the plane is demarcated by any spatial process, apart from simply moving about on it. There is therefore no need to speculate about basic human 'drives' or 'preferences' for certain space forms. Second, to show that even the most random structures are syntactically rich if we are prepared to analyse them, and we must consider that this may be a fundamental aspect of human spatial experience – to make space yield its syntactic riches by cognitively and experientially retrieving its structure – and third, and most important, because, as we hope the examples of open and closed forms show, it is the fundamental distinctions and the most elementary properties of space objects that enable us to analyse far more complex syntactic setups. We find, with Hermann Weyl that:

> What is decisive is this: the farther the analysis progresses, the more detailed the observations become and the finer the elements into which we dissect the phenomena, the simpler – and not the more complicated, as might be expected – become the basic laws, and the more completely and accurately do they explain the factual course of events.[13]

We hope to show that this is the case in the presentation of the syntax proper.

No description in the ideography has yet been offered of the ring and the closed boundary. Nor have any operations been described, apart from the most elementary operation of repetition in the minimum setup. The two questions are related since, although the existence of a neighbourhood ring may be inferred from the minimum setup, it can only arise as an independent entity if an *operation* other than repetition is performed. This is particularly clear in the case of the closed boundary (assuming we are on the surface of the earth, rather than cave dwellers). In terms of the minimum setup, the existence of a closed boundary is most improbable.

Given the ongoing process of production in the minimum setup, *an operation assigns one or more of three numbers to a configuration of objects and relations, the numbers being: one, two and many*. The numbers 'one' and 'many' are already implicit in the minimum setup in that each minimal operation adds one new object to the carrier space, and the result of such a process of repetition will be 'many' objects in the carrier space. The concepts of 'oneness' and 'manyness', both fundamental to the syntax, may therefore be said to be derivable. The concept of 'twoness' is more difficult, since it appears to be excluded by the definition of the minimum setup itself, which defined the randomness of the setup in terms of lack of a relation between two demarcations in the carrier space. The concept of 'twoness' is present in the ring but, since this is not obvious, we must take the argument a little further, then return to look at the problem of describing the ring in the ideography.

Since, as has been shown, an object placed in the carrier space has an unclear neighbourhood, it is evident that after a sufficient period of random operation of the minimum setup, a case will arise in which one object is placed sufficiently near to, or even within the neighbourhood of, another. Suppose the two objects to be *c*-objects, that is, open discs. In such a case, by following the form of reasoning which led to the identification of the neighbourhood ring of an object, there will be a subspace of the neighbourhood rings of each which will coincide. A word exists in natural language to describe such a *relation* – *between*. 'Betweenness' is a relation dependent on two objects having a spatial relation to each other. The space 'between' emerges from the relation of the two objects. Further, if the concept of 'neighbourhood' was emergent from the placing of an object in the carrier space, the concept of 'neighbour' is produced from the juxtaposition of two objects. If two objects are placed in the carrier space, sufficiently close for their neighbourhoods to overlap, then there will be a subspace of the carrier space which is common to neighbourhoods of both.

There is also a natural expression in the syntactic ideography for the relation of betweenness and, therefore, of 'twoness', viz. the expression which says that the relation of containing belongs to a pair of objects, rather than to a single object. We might provisionally write this in the form: $c_1, c_2 \supset c_3$ meaning that each object to the left of the $\supset$ sign takes part equally in the relation. Thus it can be seen that if there is exactly one term to the left of the $\supset$ sign, it gives rise to the concept of 'insideness', and if there are exactly two, to the concept of 'betweenness'.

We may now introduce into the ideography a notation for 'oneness', 'twoness' and 'manyness'. Round brackets, (), will mean one; broken brackets, ⟨⟩, two; and brace brackets, {}, many. In effect this means that two is ()(), and many might be written $\{()()\cdots()\}$. More complex syntaxes may now be constructed, and we begin by trying to solve the problem of describing a ring.

The first observation made about the ring was that, in contrast to the disc, it possessed a hole. A topologist would say that the two forms were therefore topologically nonequivalent, describing the disc as an object of genus-0, whereas the ring was an object of genus-1. The question becomes: how can this topological difference be represented syntactically, and in the syntactic ideography?

The answer lies in the *application of the ideas of twoness and betweenness to a single object*. A single *c*- or *d*-object may be 'stretched' in such a way that it *recognisably has two ends*, which would be represented in the ideography by the expression: $\langle c \rangle$. *To arrive at the ring form we add the 'betweenness' common product of this pairness in such a way as to ensure that the resultant object does not degenerate topologically to a disc again, that is, it must have a hole*. This complete transformation may thus be written $\langle c_1 \rangle \supset c_2$ (meaning: a continuous space bifurcates and is joined to itself by a further similar space, thus forming a single continuous ring), noting that it is a finite, continuous single object, and therefore $\big(\langle c_1 \rangle \supset c_2\big)$, and also that it contains a disc as its emergent product: $((c_1 \supset c_2) \supset (c_3))$ (meaning: the continuous ring contains a disc). By treating the ring itself as the object (that is, regarding the interior disc as emergent but not a participant in transformation) we may then perform the same transformation again: $\langle\langle c_4 \rangle \supset c_5 \rangle \supset c_6$ [meaning: a continuous ring bifurcates again (note not the internal disc) and is joined to itself by a similar space, thus forming a double ring] with the result that we have a continuous object with two holes. This process may be continued to create as many holes as we please, and thus raise the genus of a continuous object as high as we please. Beginning with a d-object, and performing the same transformation, we arrive at the concept of a closed boundary, or *enclosure*: $(\langle d_1 \rangle \supset d_2)$ and $\langle\langle d_3 \rangle \supset d_4 \rangle \supset d_5$.[14] In space syntax, the d version is the normal form for this object, whereas the normal form for the untransformed object is the c version. These two objects are 'boundaries' and 'spaces' respectively, or, taking into account the space inside the boundary, closed and open discs. The homologous $\rightarrow$ formula, $\rightarrow \langle ( ) \rangle \rightarrow ( ) \big)$, interprets this concept for a path, namely the path that bifurcates and meets itself, thus forming a ring.

With this language of two elementary objects, c and d, two elementary relations, $\supset$ and $\rightarrow$, and three bracketing conventions, (), ⟨⟩ and {}, the possible types of operation[15] that may be introduced into the minimum setup can be explored. The present hypothesis is that there are *eight* major types of operation, and therefore eight major types of 'syntax', all of which give rise to a principal type of settlement pattern and/or architectural complex, and whose use singly or in combination provides a method for the analysis of architectural and settlement patterns, and the morphology of building complexes. Each of these eight major syntaxes will be described in words, and expressed in the ideography in the simplest possible way (so as to make as clear as possible the relations between the syntaxes), then examined in more detail. At this stage, it is hoped that the introduction of the ideography will be justified, since it enables us both to keep an exact record of all that has been said about a pattern, but also allows us to express complexes of spatial relationships that in verbal form would be hard to follow, and even harder to formulate.

The eight major syntactic operations

Operation 1. The first syntactic operation is the one already present in the minimum setup (it will be shown shortly that this is already richer than has been described), which says, in respect of the ongoing process: *from each to the next: no relation*.

Assuming the carrier space is to the left of the leftmost relation sign (as is always the case), there is a natural expression for this operation in the ideography:

$\supset ( ) ( ) ( ) \cdots ( )$ This says that the carrier space contains an object, and another object, and another, and so on. The simplest possible expression for this operation would express this relation for the first two objects,

$\supset ( ) ( )$.

Operation 2. The second syntactic operation is based on the transformation that produces the ring or closed boundary, and it says, in respect of the ongoing process: *from each to the next: the same unitary object* – this being the effect of repeating the transformation, producing a multicellular unitary object.[16] The expression for this is,

$\supset \big(\langle\big(\langle ( ) \rangle \supset ( )\big)\rangle \supset ( )\big)$ or taking into account only the unit bracketing structures,
$\supset \big(( )\big)$,

since at each step the new object is made part of the old, with the above bracketing effect.

Operation 3. This operation combines these two ideas into that of a randomly growing but continuous aggregate, whose formal operation says in respect of the ongoing process: *from each to the next: the next is added as a neighbour to the aggregate formed by all previous objects*. This might be termed the 'pairwise growth of an aggregate', since at each step the aggregate thus far (even if it is only one object) becomes the first member of a pair, and the next object becomes the second member of the pair, these being joined in a neighbour relation onto each other. This retains the idea of randomness, while introducing the concept of an aggregate object, and pairwise could be expressed,

$\supset \langle\langle\langle ( ) ( ) \rangle ( ) \rangle ( ) \rangle$, but more simply, the concept of a randomly growing but continuous aggregate is given by

$\supset \big(( )( )\big)$.

Operation 4. The fourth syntactic operation maintains the concept of a continuous aggregate object, but instead of placing each new object in a neighbour relation to the aggregate of previous objects, it places it in a *neighbourhood* relation – that is it requires each new object to *surround* the previous. The operation therefore says in respect of the ongoing process: *from each to the next; the next is a neighbourhood of the aggregate formed by all previous objects*. The effect of this is an expanding concentric structure, which can be formally expressed as:

$$\supset \Big(\big(\langle ( ) \rangle \supset ( )\big) \supset \big(\langle ( ) \rangle \supset ( )\big)\Big), \text{ or, as simply as possible,}$$
$$\supset \big(( ) \supset ( )\big).$$

It is important to note at this stage that the second and fourth operations are unlike the first and third in that they use the concept of surrounding, which requires them to be composed of an object of the form, $(\langle()\rangle\supset())$, rather than objects of the form, (), which will serve for the other two which do not invoke the idea of surrounding. This accounts for the relatively greater complexity of the formulae for the second and fourth operations. This will shortly be seen to be one of the fundamental dimensions of variability for syntaxes.

The four operations so far described are all *local* rules which, when applied to the ongoing process of aggregation in the minimum setup, have certain *global* results (which will be discussed shortly). By local we mean that the operation works in a 'step-by-step' way, as embodied in the expression 'from each to the next'. The next four operations prescribe not a local rule, leading, under repetition, to a global outcome, but the global outcome itself, in such a way as to impose local, step-by-step order. *In other words, the first four operations control the process in local-to-global fashion, whereas the next four introduce global-to-local control.* Note that there is a relationship between operations 1 and 5, 2 and 6, 3 and 7, and 4 and 8, although this does not exhaust the interrelations.

Operation 5. The fifth operation introduces the global idea that each object produced by the ongoing process takes part in the containment of a single object. *The operation therefore says: each next object becomes part of an aggregate which contains another object*. Note that this specifies no other relation between individual aggregate members, and these can therefore be random, or emergent. Aggregate continuity is guaranteed only by the rule that relates all objects to a single object in an aggregate neighbourhood relation. Perhaps unexpectedly the introduction of a global rule has the effect of making the object finite (in all but a few variants), since the overall shape is, as it were, decided in advance, and the placing of the individual objects becomes a matter of filling available spaces from which the rule can be obeyed. To express this fully, we must say that the aggregate of objects takes the $(\langle()\rangle\supset())$ form, that is, it behaves like a ring or closed boundary,

$\supset((\langle\{()()\cdots()\}\rangle\supset())\supset())$, but at its simplest the global object is described by:
$\supset(()()\supset())$.

Operation 6. The sixth syntactic operation introduces the global idea that a single object contains the aggregate, that is a single object is the neighbourhood of an aggregate. As a neighbourhood object, the single object must have the form: $(\langle()\rangle\supset())$. The operation therefore says in respect of the ongoing process: *each next object becomes part of an aggregate which is contained by another object*. This can be expressed as

$\supset((\langle()\rangle\supset())\supset\{()()\cdots()\})$, or more simply as
$\supset(()\supset()())$.

Note that there is no relation among members of the aggregate other than being contained by the same global object.

Operation 7. The seventh operation, like the fifth, defines a global object in which a single object is to be contained by an aggregate, but it adds that one part of the aggregate is containing another part, and the single object contained is between them, that is, contained by both. This has the effect of making the single object into an enclosed ring (if the fifth operation generates 'plazas', the seventh initially generates 'ring-streets', an equally pervasive form). The operation therefore says: *each next object becomes part of one of two subaggregates, one of which contains the other, and which between them contain a single object*. This may be expressed as

$\supset(\langle\{()()\cdots()\}\supset\{()()\cdots()\}\rangle\supset())$, or more simply as
$\supset(\langle()()\supset()()\rangle\supset())$.

Operation 8. The eighth operation reverses the seventh in much the same way as the sixth reverses the fifth: two single objects, one inside the other, have between them an aggregate. This means that the aggregate unfolds itself within two objects of the form $(\langle()\rangle\supset())$, one of which is inside the other. The operation therefore reads: *each next object becomes part of an aggregate which* is *contained between two objects, one of which contains the other*. This may be written:

$\supset(\langle(\langle()\rangle\supset())\supset(\langle()\rangle\supset())\rangle\supset\{()()\cdots()\})$, or more simply as
$\supset(\langle()\supset()\rangle\supset()())$.

From operation 1 to operation 8 the increasing complexity of syntax formulae expresses the increasing degree to which the 'global' structure prevails over the 'local', with the highest numbers having the most global structure. Therefore it is reasonable to say that patterns built from strong global operations have 'more' order in them than 'locally dominated' syntaxes. This does not mean that 'local' syntaxes do not have global structure; they have very strong global structure, but it is a global structure that emerges from the ongoing process, under the influence of its rule. These emergent patterns will shortly be examined in more detail.

Inspection of the eight formulae reveals another principal dimension of variability which occurs between the odd and even numbered syntaxes. If we begin with an initial object and consider the way a particular syntactic operation builds on it, it can be seen that even-numbered syntaxes, after 2, build control to the left of the initial object (as though the initial object were the linguistic 'object' of the formal 'sentence', and the left-added structure were the 'subject' of the 'sentence'), which always has the 'surrounding' form $(()\supset())$, taking the form of a single object. In contrast, odd-numbered syntaxes build to the right of the initial object, accumulate, and allow the accumulation to control whatever is built to the right (as though the initial objects were the 'subjects' of the 'sentence', and the added structure was the 'object'). This is the formal difference between *distributed* and *nondistributed* syntaxes. In distributed syntaxes, the odd-numbered syntaxes, every primary object plays an equal part in constructing the global pattern (whether this is globally defined or emerges from a local operation), and the aggregate therefore predominates. In nondistributed syntaxes, the even-numbered syntaxes, there are always one or more unitary loci of control externally placed on the aggregate, which impose a superordinate control on it, and therefore the unitary object predominates. To use a contemporary illustration of this fundamental distinction, a street-pattern is a global distributed syntax, whereas a high-rise estate is a global nondistributed syntax. The shift from distributed to nondistributed syntaxes is one of the fundamental dimensions of the shift to urban space patterns in the past century or so.

It is hypothesised that these general dimensions of variability of syntaxes (local-to-global and global-to-local, distributed and nondistributed) can be well ordered in terms of the model presented in Figure 6.5, which shows the interrelationships of all the syntaxes by showing how they are derived from one another. The model initially postulates that operation 2, $(()\supset())$, is the 'prime' of operation 1, and from there operation 1 is combined with operation 2 to make syntax 3, then 3 is primed to make syntax 4, then 4 is combined with 1 to make

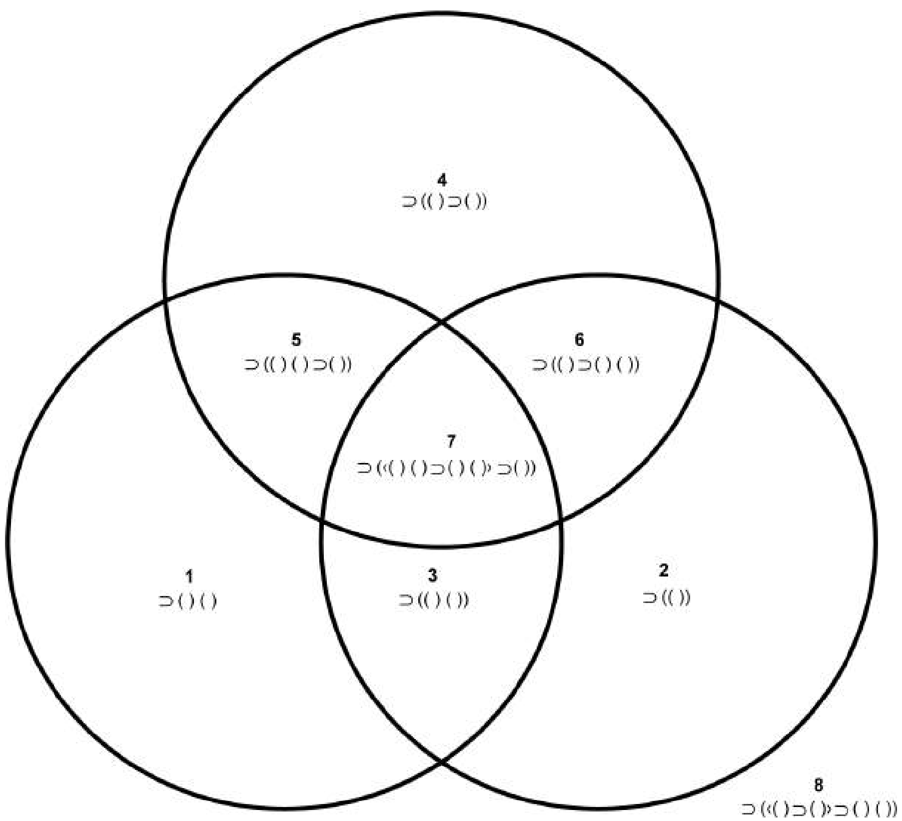


Figure 6.5 Evolutionary relations of syntactic rules.

operation 5 and with 2 to make operation 6, then with 3 to make operation 7, and then operation 7 is primed to make operation 8.

Alternatively the eight syntaxes may be seen as the unfolding of the following rule for forming syntax rules (or 'rule-rule'), imposed on the minimum setup:

(a) if ()() then add as many as you like (that is, if two objects are juxtaposed in a formula without a relation between them, then the two can become many);
(b) if () ⊃ then () remains one (that is, if a relation sign follows one object in a formula, then it remains as one);
(c) if () ⊃ then () is $(()\supset())$ (that is, if an object contains, then it is a ring).

In other words, given the generic difference between syntaxes 1 and 2, we have simply explored the possibility of shifting brackets relative to relation signs, until all more complex versions that can be unfolded on the carrier space must be combinations of these. The eight are the fundamentally different types of syntax.

In order to embody these abstractly stated operations in rule structures in the full sense, so that they can be imposed on the minimum setup to give real space patterns, which may then be compared to reality and made the subject of a socio-spatial theory, we need to build objects into the operation formulae. This can be done by means of two simple additional rules. The first is justified by the observation that the normal form of an independent structure of the form $(()\supset())$ is $(\langle d\rangle\supset$挃$)$, that is, the enclosure that makes the closed disc. Since we have already defined this form as the 'prime' of the open disc and assigned it to the family of nondistributed syntaxes, we may also regard $(\langle d\rangle\supset d)$ as the elementary nondistributed, or 'primed', object. Conversely an open disc is the elementary distributed, or 'unprimed', object. (See Figure 6.5.)

The second follows from the definition of 'open' and 'closed' objects given earlier; namely, that an object is closed if it has its own boundary. From this we say that an object will be closed if it takes part in containing. Note that these are not strict rules but conventions which give rise to the normal, real space forms of the syntax. In fact, the exceptions to these conventions are often interesting and informative.

Syntactical analysis applied to some existing forms

We may now look more carefully at the syntaxes,[17] and offer a new way of approaching a number of classic problems in the morphological study of settlement patterns, and the evolution of architectural and urban form. In this discussion, reference will not be made to social and organisational variables, since these are dealt with as a whole in section 4, which outlines a general theory of the social formation of space patterns.

Each of the eight syntaxes gives rise to an elementary object, namely that constituted by the least possible interpretation of its operation, with regard for the

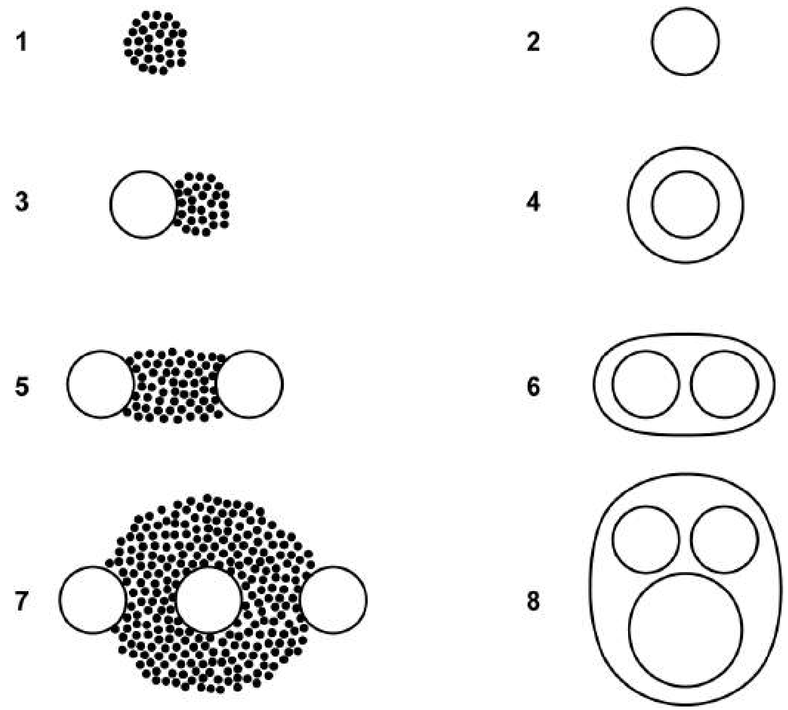


Figure 6.6 Elementary syntactic objects.

relations between open and closed objects, and distributed and nondistributed syntaxes. These are given in Figure 6.6.

It may or may not be noteworthy that if each space and each boundary are counted as an object, then each 'least object' has exactly the number of objects of its syntax number. This may be fortuitous, or even contrived, or it may be simply the general result of a 2-object being more complex than a 1-object, a 4-object being more complex than a 3-object, and so on.

The 'least' object is the natural starting point of a syntactic process and gives its dominant morphological realisations in most cases. On the other hand, it is not a necessary starting point. Much of the variability of syntactic forms comes from variation in the 'least objects' of which the global pattern is constituted. For example, a global 5-form may have 4-objects as its constituents, and so on. In particular, most examples of 1-syntactic settlement patterns have fairly complex objects fed into the 1-syntactic process. The arguments set out in section 4 of this paper suggest why this should be so.

The most important product of each syntax is the global pattern produced when the syntax is defined on the minimum setup for a sufficient period for emergent patterns to appear. This is particularly important for the distributed syntaxes, beginning with the 1-syntax, which has not quite yet extracted all the riches from the minimum setup.

When the minimum setup was first defined, it was not specified whether or not the carrier space should be regarded as being bounded. In taking the surface of the globe, this problem was avoided. But this is a little unrealistic. The 1-syntax defines its own carrier space in the following way. The first object is placed at random, and then another an arbitrary distance away, in an arbitrary direction, possibly following topographical or resource constraints. By this time it is possible for the third object to treat the zone within which the first pair can be thought to lie as the carrier space to which it will relate. As the process develops, each object as it is placed will either be surrounded at some distance by other objects, in which case it will not be near the edge of the evolving carrier space, or it will be only partly surrounded by other objects, in which case it will be near the edge, or even outside. If this latter case arises, then the next object will be placed back in the region of the other objects in order to follow the rule of belonging to the carrier space. Thus the 1-syntax adds a small measure of structure to the minimum setup, in the form of a local rule comparable to that which, Thom argued,[18] maintains the spatial coherence of a cloud of mosquitoes – that if each mosquito, while moving randomly with respect to each other mosquito, sees half his field of vision free of mosquitoes, he moves in the direction of mosquitoes. This is, we believe, the minimum syntactic rule for a spatially coherent aggregate, and *its global result is that, through the distributed repetition of its local rule it defines the carrier space*. Thus there is a means of producing global differentiation in the man-made landscape without the invocation of boundaries, and even without a clear idea of the limits of the carrier space of a particular collection (that is, members of a society who place objects on the land). Thus in a 1-syntax, the 'smallest object' is what was earlier defined as a *differentiated space* – an open disc which is recognisable by virtue of being differentiated rather than distinguished by a boundary – and the global structure that results is also a differentiated space – the *large open disc that now constitutes the carrier space*.

There is also a 1-syntactic path-structure (all syntaxes have space forms and path forms) which arises by an exactly analogous process. As each object is added it carries a path from the last, thus generating a *carrier-path* analogous to the carrier space. As the surface becomes dense the paths inevitably form nets, and the net is the global path-syntax associated with the 1-syntax. Its simplest interpretation is 'being strung along a path', which is one of the most important, albeit minimal, determinants of settlement patterns. The capability of the model to describe such minimally ordered space and path setups in a unified theory is, we hope, one of the justifications for the rather elaborate construction of the concept of a morphic language.

A settlement pattern, which we would now call 1-syntactic, has been described by the anthropologist Fortes. Writing of the Tallensi, a tribal society of considerable intricacy, with highly structured households, Fortes observed that the landscape of Taleland (the global, open disc, boundaryless carrier space) had the following appearance:

> From the top of the Tong hills, looking northwards, one has a view of what seems to be an endless plain, dotted meagrely with trees and studded closely, as far as the eye can reach, with homesteads. They are identical in appearance, squat, circular, drab-grey or red, like the soil itself, mostly thatch roofed, and seem to be scattered indiscriminately, some close together, others further apart. There is nothing to indicate where one settlement begins and another ends.[19]

Figure 6.7 shows a segment of the pattern.[20] The settlement structure of the Tallensi also includes other randomly placed spaces, not shown on the illustration, which are their 'sacred places' or 'earth-shrines'. These are not enclosures but 'groves of trees, a pool or stream, a pile of boulders, a single prominent tree, or simply a bare patch in the fields'. This is a fair sample of the possible methods for demarcating the landscape without boundaries. It shows the Tallensi are being consistent. For their sacred spaces they use 1-objects; for their settlement patterns the 1-syntax.[21]

The 2-syntax also features prominently in the record of the evolution of settlement patterns. A very pure example is the multicellular 'kiva-block' at Kaituthlanna,[22] which appears as an antecedent to the classic ruined pueblos of

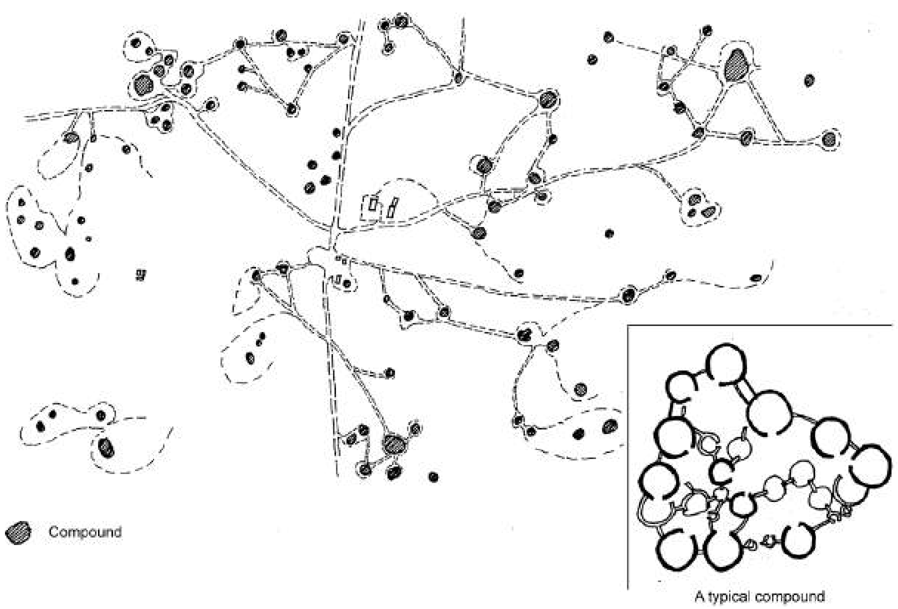


Figure 6.7 Tongo – a Tallensi settlement (after Prussin, 1969, pp. 56, 59).

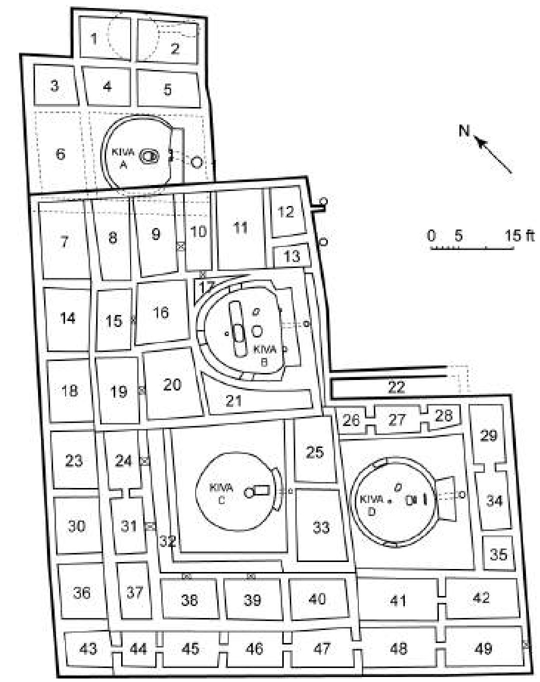


Figure 6.8 Pueblo ruin at Kiatuthlanna (after Roberts, 1931, pp. 91–92).

the American Southwest. Figure 6.8 shows the characteristic form, lacking internal connection between the separate cells (these arise later). The large cell is the kiva, which was probably the original built form; the multicellular structure is, as it were, 'unfolded' from the kiva outwards. Again there is syntactic consistency between the use of the simple 2-object for the sacred space, and the 2-syntax for the settlement model. The ethnographer from whose work this illustration is taken, having taken for granted that the structures must be 'defensive', commented that this could hardly be the case since they could easily be attacked from above, there being an overhanging cliff available. This is a typical example of the failure of simple 'causal' explanation of built forms and settlement patterns, and their rejection by many ethnographers. The 2-path syntax associated with the 2-space syntax is a series of path segments leaving the carrier path at intervals, and each arriving at a limit, or dead end.

Back-to-back housing is the nearest modem version of this morphology, although early examples including possibly Çatal Hüyük,[23] thought to be the world's earliest town, appear to have an analogous form.

The 3-syntax is one of the most widely used forms of settlement organisation, and examples, varying in density and small object composition, have been found in the ethnographic records of many parts of the world. It is still a dominant urban

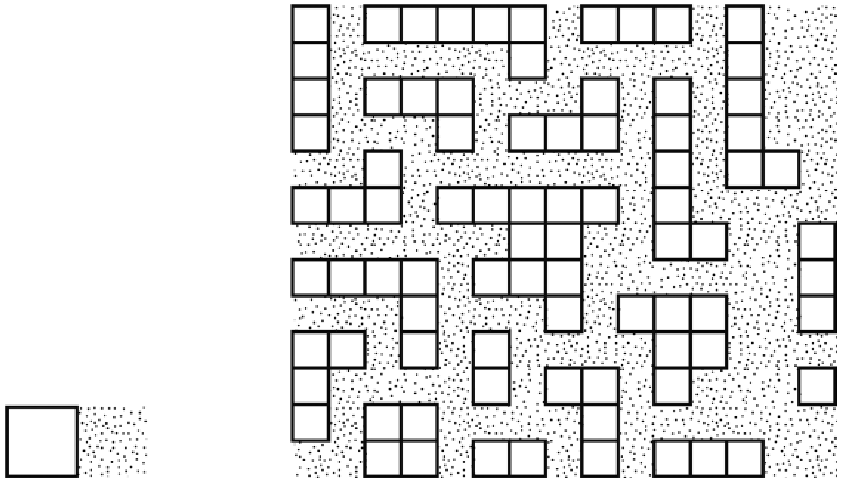


Figure 6.9 (a) 3-object and (b) randomised global pattern.

pattern in certain parts of the world and, for example, survives in parts of the City of London. The most initially surprising aspect of the 3-syntax is its global form, or rather its family of global forms. These can be discovered by means of simple experiments, either with pieces of black and white card, or even with pen and paper.

If the setup has for its object the 3-object (Figure 6.9a), and a 3-syntactic rule (continuous random aggregation) is applied, requiring objects to be joined facewise (there is a universal syntactic rule forbidding the joining of rectilinear forms by vertices) by the open discs (if we require both components of the 3-object to be facewise joined, then the only result will be a linear string – this occurs frequently, but is not our immediate object), then provided the process is properly randomised the following type of global pattern will begin to emerge (Figure 6.9b).

This is an oversymmetrical version of the global syntax, since the process was carried out as though on a grid to show the rule working. The defining properties of the global pattern are that

(a) a continuous structure of asymmetrical open space is unfolded, varying in width, and containing at odd intervals what may be called 'beads' – in fact, this form is known to us as the 'string of open beads';
(b) this open space structure eventually forms rings, so that there are always 'two ways round' from any point to any other point – it becomes 'rings of open beads' (unless there are additional constraints that force it to unfold linearly and remain 'strings');
(c) the closed discs of the 3-object are divided into islands or clumps by the formation of the rings; with the result that
(d) every closed disc (in effect the dwelling) is directly attached to a potentially infinite, and intricate, open space structure, and will almost certainly be near a 'bead' of some kind, that is, a larger piece of open space;
(e) the 'rings' thus formed are both open and enclosed in the senses already defined.

It should be noted that apart from these strong structural properties, the asymmetry of the pattern is a syntactic product, a result of the inherent improbability of geometrical order in the minimum setup.

This settlement form, which will be the subject of a separate monograph, has often been noted by writers on settlement morphology, but usually classified either as 'disordered', or simply as 'nucleated', 'clustered', 'organic' or some other somewhat unhelpful term. The syntactic analysis reveals the profound spatial ordering of such settlements, a fact easily appreciated by the average tourist, if not by the average morphologist. Figures 6.10 and 6.11, Muker and Hawes respectively,[24] both in the Yorkshire dales, show a very small settlement and a small English town at different levels in the unfolding of a 3-syntax.

As frequently happens, in Hawes the back areas of the 'island' aggregates become closed off, but the general syntactic form remains. Here there is both a 'string of open beads' (the narrow part) and 'rings of open beads' (the broader section growing randomly, and therefore in all directions). The path syntax associated with this pattern is one in which a path segment leaves the carrier path, and eventually returns to it by another route: the nonlimit sequence. There is a sequence of traversals, but no dead ends.

In contrast to the 'profanity' of the 3-syntax, the 4-syntax (with its associated path structure, the 'limit sequence', that is a sequence leading to a 'deepest space') is strongly associated with the sacred. Sacred buildings, from English churches to

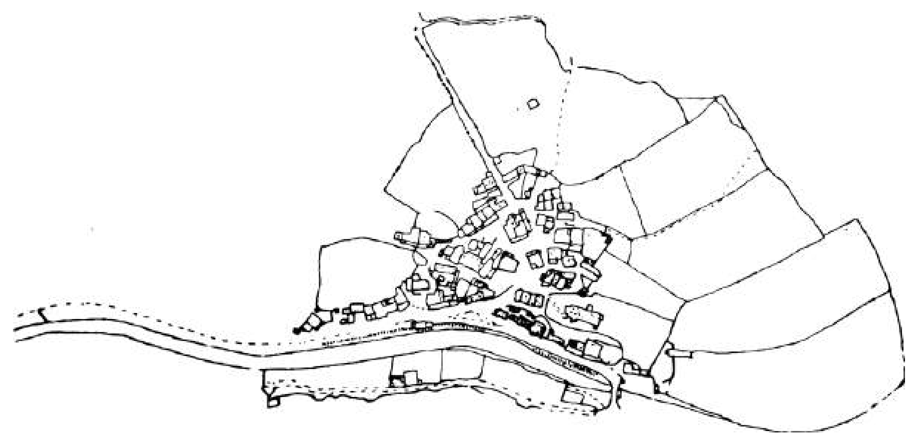


Figure 6.10 Muker (after OSC, 1888–1893).

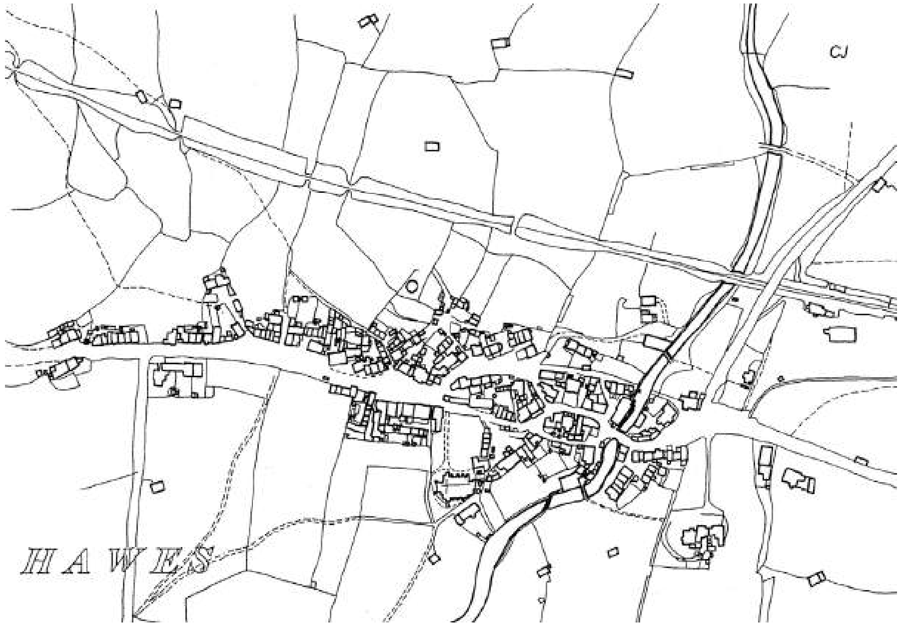


Figure 6.11 Hawes (after OSC, 1888–1893).

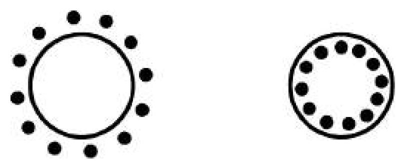


Figure 6.12 (a) Closed disc with an 'open space barrier' around itself; (b) Closed disc with a 'cloister morphology'.

the Summer Palace in Beijing, have concentric overall morphologies, and explore the path concept of a limit sequence. Most interesting are the variants of the 4-syntax. For example, there are a pair of forms which use the 4-rule to make a concentric object, but make it from an open and a closed disc rather than from two closed discs. One is the closed disc which defines an 'open space barrier' around itself – that is the open contains the closed, the other is the 'cloister' morphology, in which the closed contains the open (Figure 6.12).

Suburban morphologies are usually based on 4-objects. The open type of American suburb is based on the 'open space barrier' morphology; the British suburb on the basic double closed disc. In the more compressed urban setting, the sequence is substituted for the full concentric form.

Thus each of the first four syntaxes leads to emergent structure, that is, more structure than was specified in the syntax rule. What is less obvious is that each

defines as its emergent structure one of the four basic objects of the syntax. The 1-syntax generates an open disc; the 2-syntax the closed disc; the 3-syntax the open ring; and the 4-syntax the closed ring. We are far from clear why this should be so, nor are we clear if it is significant. It does happen, however, that, in the same order, the emergent forms become the defining properties of the next four syntaxes. We thus move on from the 'step-by-step', or 'local-to-global' cases, in which these structures emerge *from* an aggregate, to the 'global-to-local' cases where the structures are defined *for* the aggregate.

The 5-syntax generates a family of forms that are both well-known and well recognised, possibly because they are often geometrically fairly obvious. Also, because the form is a simple one and also finite, it occurs at many levels. At the simplest level, the 5-syntactic defining property of an aggregate containing (that is constructing) an open disc generates the simple courtyard form in its various realisations, including the version arising from a bifurcated, continuous aggregate that meets itself – as in Figure 6.13a and the double pair form, which historically appears to grow into the example above, and may be logically prior to it – as in Figure 6.13b.

On a larger scale, it generates the 'green' village and the 'plaza' pueblo. Because they are finite, 'plaza' forms rarely produce towns in the proper sense, but there is a variation on the 5-syntax which does generate towns, because it is the only nonfinite version of the rule. This is where the aggregate breaks into two, and the open disc is then enclosed *between* the two aggregates and can therefore continue to grow and expand at either end. This is the classic 'long-wide street' morphology, in which the urban form is little more than a single, wide, often long, 'street'. In the many examples of this syntax among English settlements, it is frequently the case that the elongated space is wider in the middle than at the ends, thus emphasising the 5-syntactic form. More interesting, however, is the fact that in a high proportion of such cases the mode of adding further objects is in strips or pairs of strips away from the 'street', with an opening directly onto the street. This is essentially a repetition on a small scale of the same syntactic principle. It expands the global open space into a series of long 'fingers' reaching deep into the areas behind the objects fronting on the main street. In this morphology, rings very rarely occur. Bedale, shown in Figure 6.14,[25] while it does possess a single ring, is a good example of this morphology.

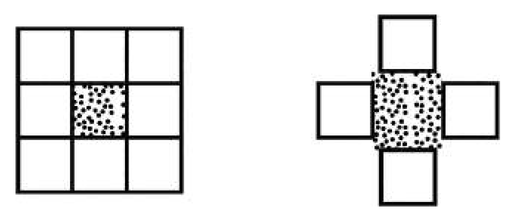


Figure 6.13 Courtyard forms: (a) a continuous aggregate that meets itself; (b) a double-pair form that grows into the form of (a).

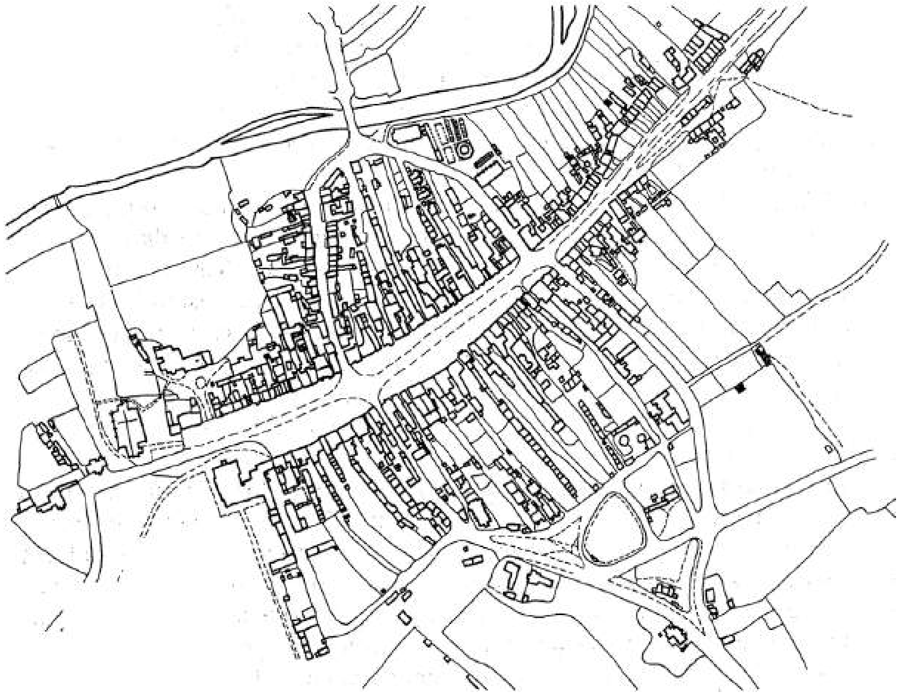


Figure 6.14 Bedale (after OSC, 1888–1893).

The path morphology associated with the 5-syntax is produced by the idea of an aggregate controlling permeability to itself through its enclosed open disc. This has the characteristic form of a 'star'.

Whereas the 5-syntax 'glues together' an aggregate of closed discs by means of an enclosed open disc, the 6-syntax 'binds together' an aggregate of closed discs by means of a closed disc which encloses the aggregate. Among the characteristic morphologies produced by this rule are the modern 'block' and the 'estate'. The associated path morphology is the 'tree' of nontraversing paths (that is a path segment leading to a set of branching segments which themselves have branches, but which are eventually limits) to which in certain realisations we assign the name 'corridors'. Since these relations are far from obvious (6-syntaxes are much less well-recognised morphologies than 5-syntaxes) they need to be examined with some care.

First, it should be recalled that the 2-syntactic object was a closed boundary, not a closed disc. The closed disc emerged from the fact that if there had been nothing contained within the closed boundary, it would have degenerated to a simple d-object, and no transformation would have taken place. Similarly for the 4-syntax, the second closed boundary contained a ring since, had it not done so and been pressed instead directly onto the closed disc, the transformation would have

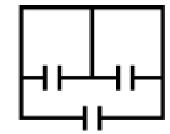


Figure 6.15 A large boundary containing a pair of closed discs – rooms – against which it is directly pressed on three sides.

degenerated into a closed disc again. At the 6-syntactic level, these consequences of pressing a boundary directly onto another no longer arise, since the second boundary is now to enclose an aggregate, one moreover that lacks other internal relations to 'glue' it together.

The second boundary may therefore in principle be pressed directly onto the boundaries of the closed discs it contains. A simple object illustrates this in Figure 6.15.

This is unlike the 2-syntactic multicellular object in that it has an extra structure defined on it. This extra structure produces the 6-syntactic path morphology, in that a nontraversing path controls permeability to an aggregate (two in this case) of closed discs, giving the elementary branching structure of the 'tree' path form. Less obviously, but we hope no less necessarily, the two closed discs should be seen as being bound together by a complete outer boundary which, in part, is pressed directly onto the boundaries of the closed disc. Formally speaking, this can be justified by a perfect homology of space and path formulae:

$$(\langle d \rangle) \supset d) \rightarrow \{()()\ldots()\},$$

and

$$(\langle d \rangle) \supset d) \supset \{()()\ldots()\}.$$

Intuitively it can be argued that the lack of internal space structure forces a separate – and therefore nontraversing – path system into existence in order to make the structure permeable. A clear illustration of a 6-syntactic morphology, where an aggregate of closed discs (they have in fact additional internal structure, but this does not concern us in analysing the global morphology) is tightly bound together by a closed boundary and is internally connected by a (slightly imperfect) tree of nontraversing paths, is the 'workers' quarter' in El Amarna in ancient Egypt, shown in Figure 6.16.[26]

The characteristic forms of 'estate' built increasingly over the past century offer some of the clearest examples of 6-syntactic morphologies, often combining the 6-syntactic version of the block, with the concept of an 'estate', which is a name we give to the closed disc formed by the estate boundary and the space in which the

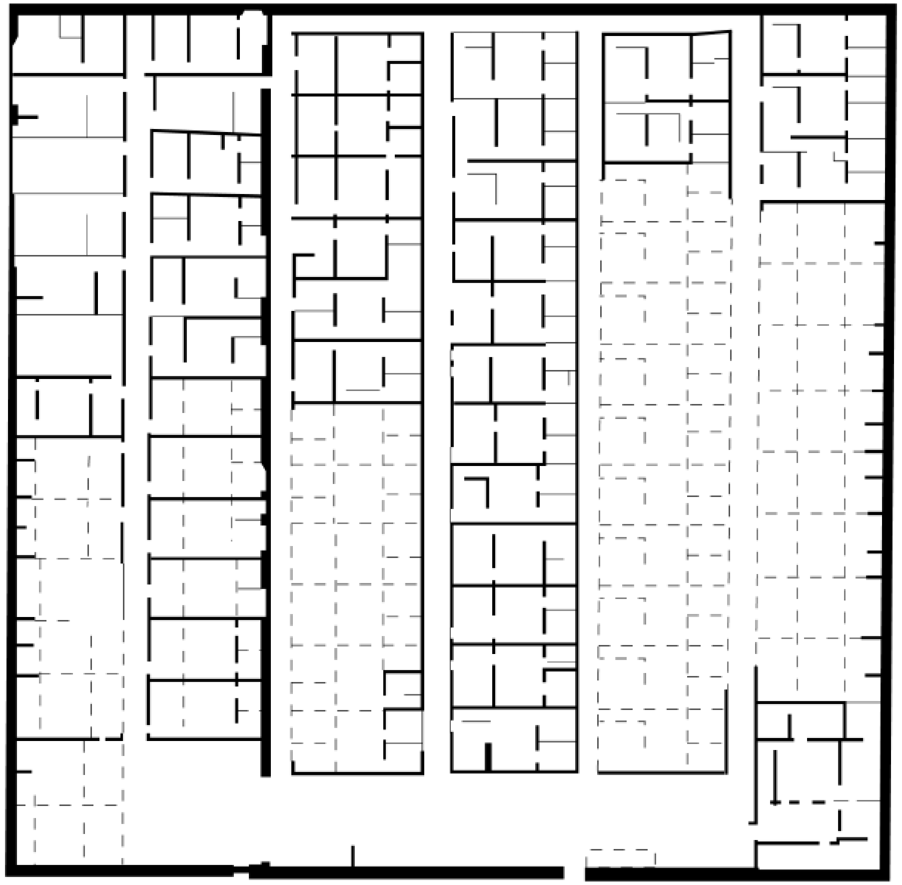


Figure 6.16 El Amarna, workmen's village (after Peet et al, 1923).

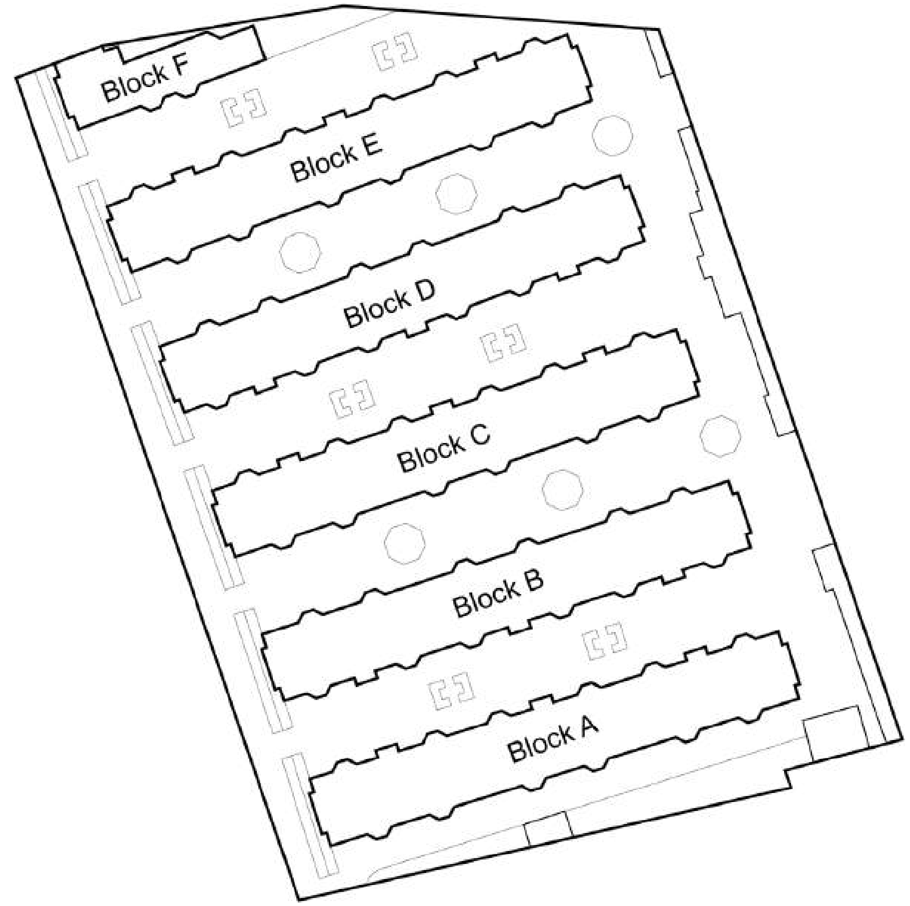


Figure 6.17 Samuel Lewis Trust dwellings (OSC, 1955).

'blocks' are so carefully 'laid out'. It is clear that no arrangements that leave the two levels of 6-syntax invariant in such a morphology will have any effect on the final syntax of the result. A characteristic example is given in Figure 6.17.[27]

The defining property of the 7-syntax is the open, enclosed ring (and its associated path morphology, the traversing ring-path), that is, a ring which is enclosed not by its own boundary but by an aggregation of closed discs whose boundaries all belong to them. This requires a split in the aggregate such that one subaggregate is inside the other, and the ring is therefore defined by being between the two. The elementary object already defined interprets this by putting one closed disc between two others, with an open ring surrounding the middle disc. A more developed and lifelike realisation would be as in the diagram in Figure 6.18.

In its elementary unfoldings, the 7-syntax gives rise to the classic, minimal '*ring-street*' morphology, which has been a generic settlement form in many parts of the world. Two slightly differing examples are given, one from southern France, the

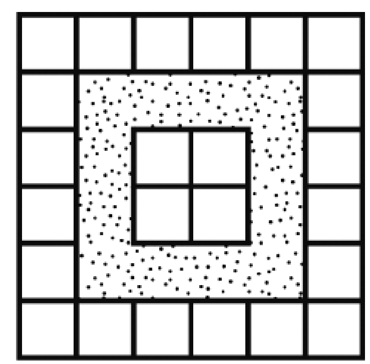


Figure 6.18 Derivation of a street pattern from the 7-formula.

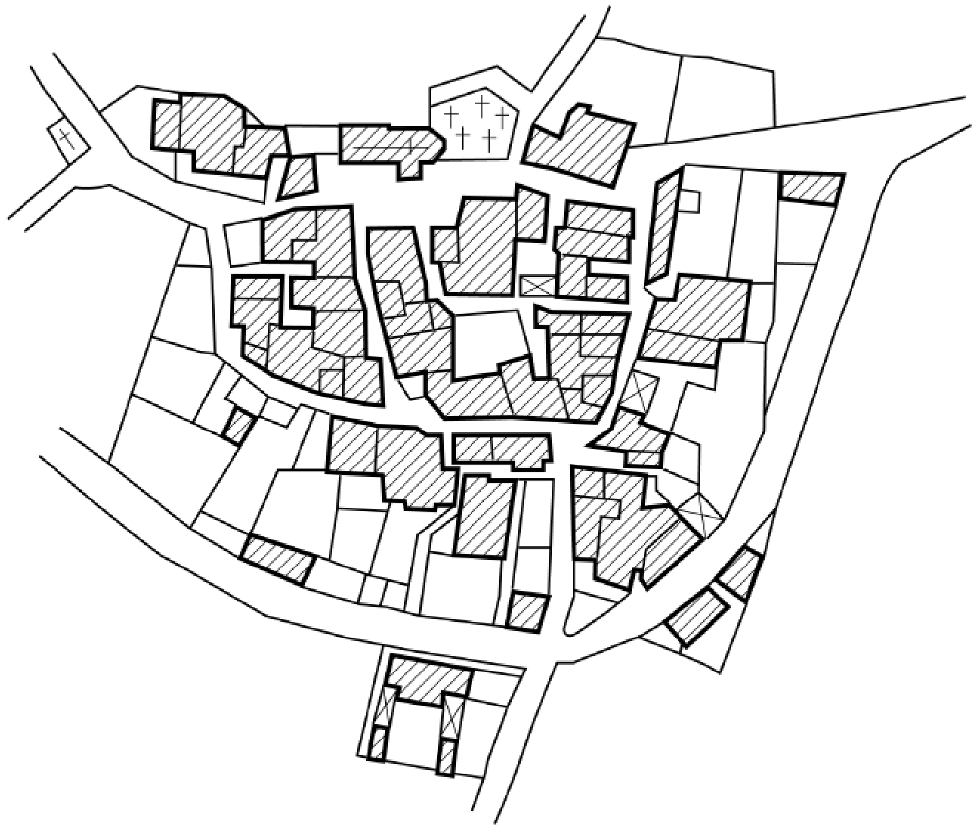


Figure 6.19 Village of Saint-Trinit (after SD, 1934).

other from northern England, the villages of Saint-Trinit and Burton in Lonsdale, Figures 6.19 and 6.20, respectively.

The most important morphological result of the 7-syntax is, however, the concept of a '*street pattern*' strongly separating front and back (unlike the 3-syntax), which is exactly given by the continuous aggregation of open enclosed rings. If we take, for example, the case of the somewhat idealised French medieval town (nevertheless a real one), see Figure 6.21, it can be seen that its global description is of a ring of open enclosed rings containing an open disc. The way in which the open rings were first constituted (prior to infill of any 'holes') is shown in a plan of the town of medieval Conwy, Figure 6.22.[28]

Thus it can be seen that from the foundation, a 'street', if it is not the long wide 5-syntactic street form, is only constituted by being part of a street pattern, based on at least one open enclosed ring. If this analysis is correct, then it would appear that recent attempts by designers to recreate 'the street' by means of wider 'access decks' do violence to the syntactic nature of the street form. It would appear that the lay rejection of such improvisations is morphologically correct.

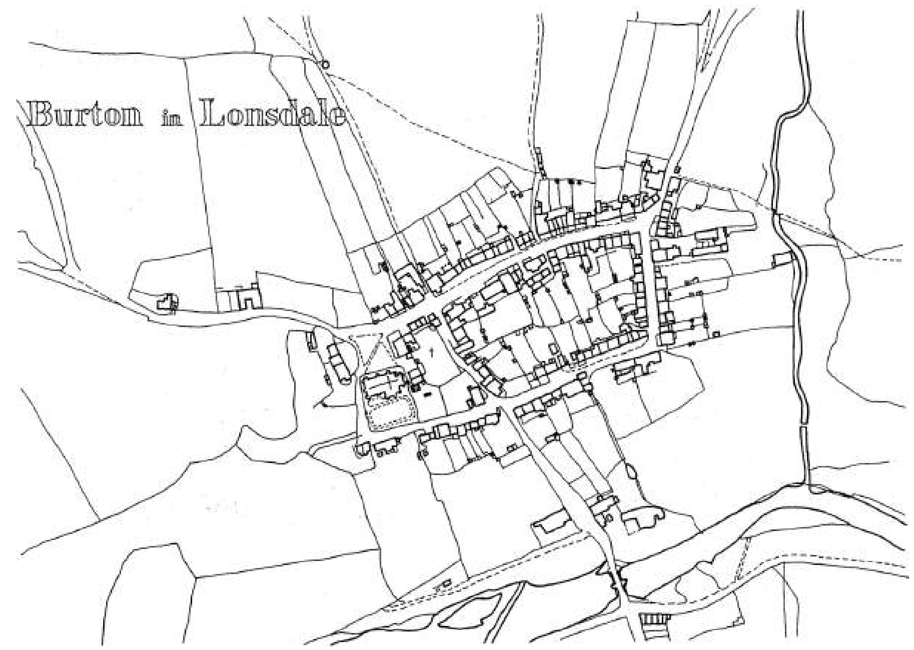


Figure 6.20 Burton in Lonsdale (after OSC, 1883–1893).

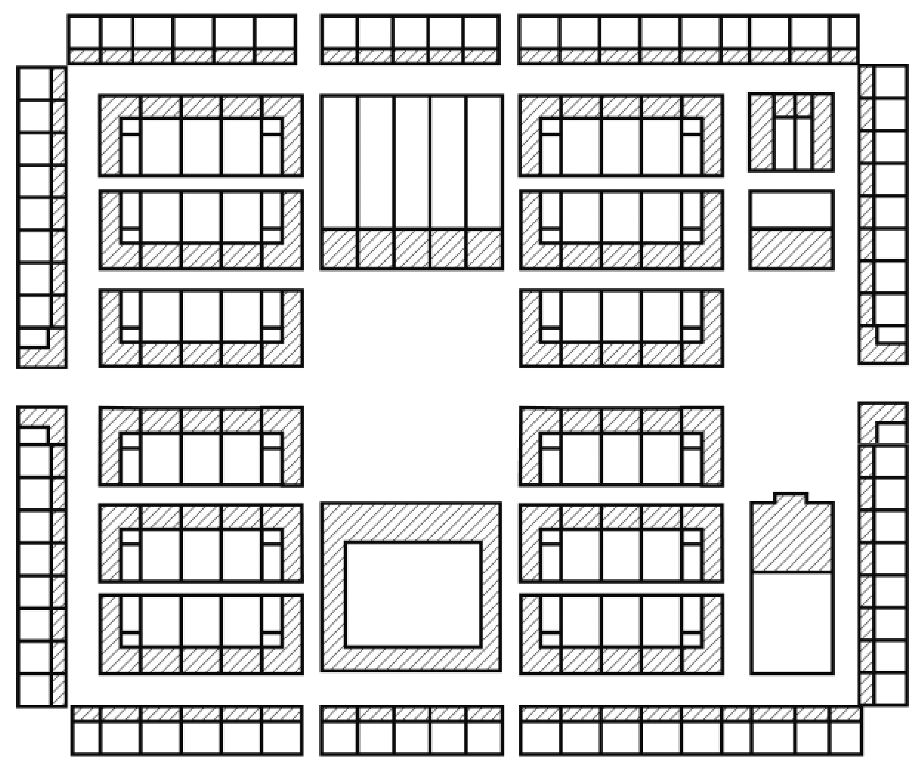


Figure 6.21 The New Town of Erlangen, founded by the Huguenots (after Gutkind, 1964, p. 222).

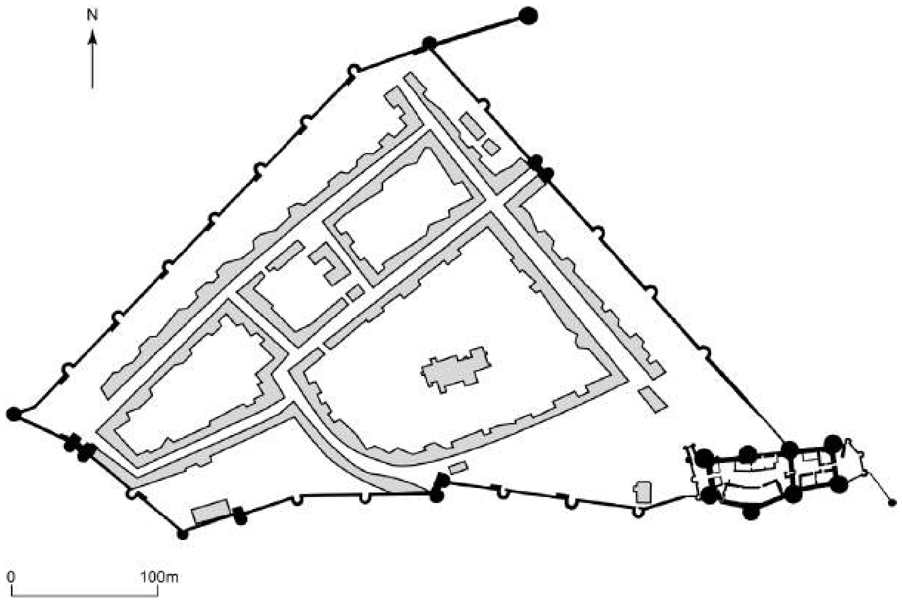


Figure 6.22 Medieval Conwy (after HMSO, 1957)

There is another important variant of the 7-syntax, whose identification solves an outstanding problem in the study of settlement evolution. This problem is to explain why the 'Pueblo Indians' stopped building 'plaza' settlements and began to build settlements consisting of compact lines of dwellings. Reed commented:

> Interpretation of the changes during the last few centuries in the Upper Rio Grande, from the front directed Anasazi plan to the hollow square layout (ubiquitous, apparently, during Pueblo IV) to predominance of parallel alignments, except among the Tewa, is beyond me ... why the linear parallel alignments supersede the unified hollow square in the west, I have no idea.[29]

The answer is that just as there is a linear version of the 5-syntax, which works by interpreting containment in terms of 'betweenness' rather than 'insideness' (the long wide street form), so there is exactly such a version of the 7-syntax, but involving three lines of aggregated closed discs rather than two. This is easily understood if we return to the 'least object' for 7, which can clearly develop either way. The formula for the 'three line' unfolding is as follows:

$\supset(\{()()\cdots()\}\{()()\cdots()\}\supset\{()()\cdots()\}\supset())$, which can be read: a pair of continuous subaggregates have between them a continuous subaggregate, and

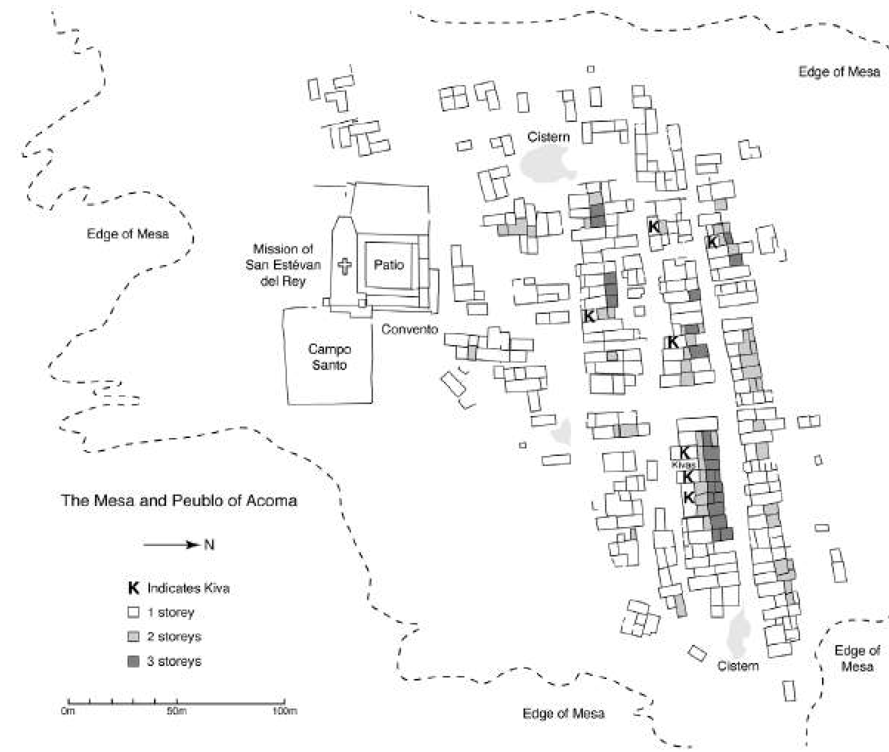


Figure 6.23 Acoma (after Stubbs 1950, Figure 20).

the pair formed by the pair of subaggregates and a single subaggregate contains a continuous space (see Figure 6.23).[30] Since neither defence nor climate appear to have any role whatsoever in this transformation, might it not be sufficient to assign syntactic reasons for the change?

The 7-syntax is the most global distributed syntax (that is, each primary cell equally constitutes the predetermined global morphology). The reason why urban form has not progressed beyond the 7-syntactic form but retreated from it is, of course, that there is nothing beyond it. It is all that is possible.

The 8-syntax is the most global nondistributed syntax, and its unfolding locates all primary cells within the inner and outer boundaries of a closed ring, without further internal relations, as in the 6-syntax. An idealised version of a typical object might have the form seen in Figure 6.24.

Although interesting examples of this syntax exist in ethnographic records, including, for example, the 'single building settlements form' of the Hakka tribe in China (Figure 6.25), and even the mile-wide 'great kraal' of Shaka the Zulu king (Figure 6.26), the most spectacular realisations are modern, or recent.

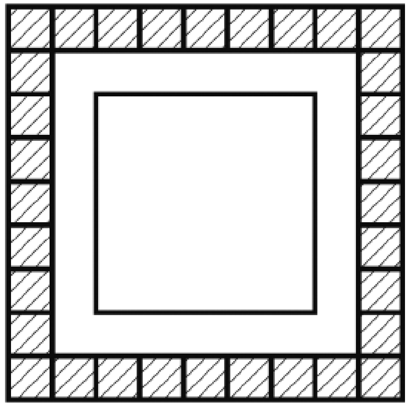


Figure 6.24 Idealised form of 8-syntax.

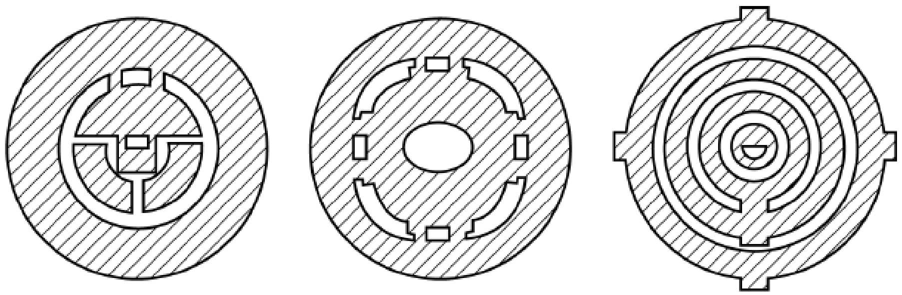


Figure 6.25 Circular plan types of Hakka dwellings (after Boyd 1962, p. 105).

For example, the classic models for prisons in the 'panopticon' era moved from 6-syntaxes to elaborations on the 8-syntax. Take, for example, the design of Bevans,[31] Figure 6.27, which can be described as: a ring of closed rings containing aggregates of closed discs, contains a closed disc. It turns out to be an exact inversion of the 'urban' distributed form of the New Town of Erlangen shown in Figure 6.21, above. The path morphology associated with the 8-syntax is the nontraversing ring path.

Figure 6.28 summarises the eight syntaxes and their principal morphological realisations.

From each syntactic generator, a family of related forms can be generated by introducing further bracketing into a formula, while leaving the *defining relation* of the syntax invariant. For example (simplifying formulae for the sake of clarity), Figure 6.29 shows an example of the variants in the 5-syntax:

From these it can be seen that an overall rebracketing keeps the form relatively localised, while the 'left only' bracketing requires more closed discs to be related to the open discs, and this increases the degree of global order.

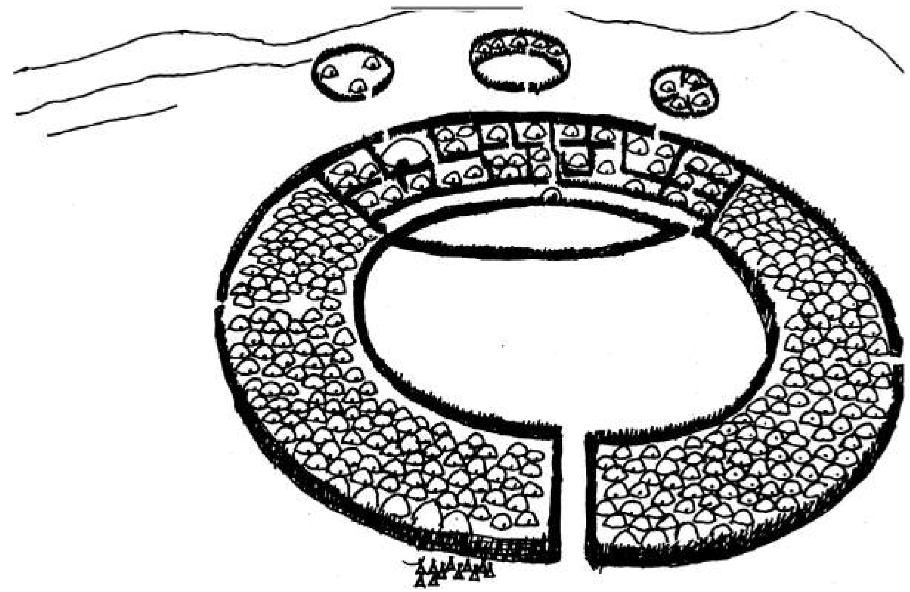


Figure 6.26 Zulu royal kraal (after Gluckman 1960, p. 157).

The derivation of a street pattern from the 7-formula follows from a development of the bracketing structure. If we begin from the form $\{\,\}\supset\{\,\}\supset(\,)$ (as given in the 7-syntactic object shown in Figure 6.18 earlier), it is not easy to see how the double containing relation (the open ring is between two subaggregates, one of which is inside the other) is retained in a street pattern based on a set of open rings. The growth process is as follows: at least the second of the pairs of aggregates becomes a pair, $\{\,\}\supset\{\,\}\supset(\,)$. (If the first aggregate also becomes a pair, or even a pair of pairs, the essential transformation remains.) The form will then, of necessity, contain a pair of intersecting open rings as the unitary structure on the right of the formula; see Figure 6.30.

Two clear realisations of this unfolding are seventeenth-century Peterborough and Hertford (Figures 6.31 and 6.32[32]). It can be seen that this procedure may be followed to generate any number of intersecting open rings.

Some interesting observations on the syntactic nature of streets can be offered at this point. It can be seen that although a street is only a street by virtue of its membership of a street pattern with the minimum form of a simple open ring, at the same time each street in a sufficiently rich setup is itself the unique intersection of a pair of open rings. This means that each ring differentiates naturally into four segments with a unique set of local syntactic relations. This seems a reasonable formal approximation of an essential intuitive property of a street: that it is uniquely differentiated syntactically while being continuously connected to a continuous structure of open space. Thus a street is not simply a certain kind of enclosure. It is

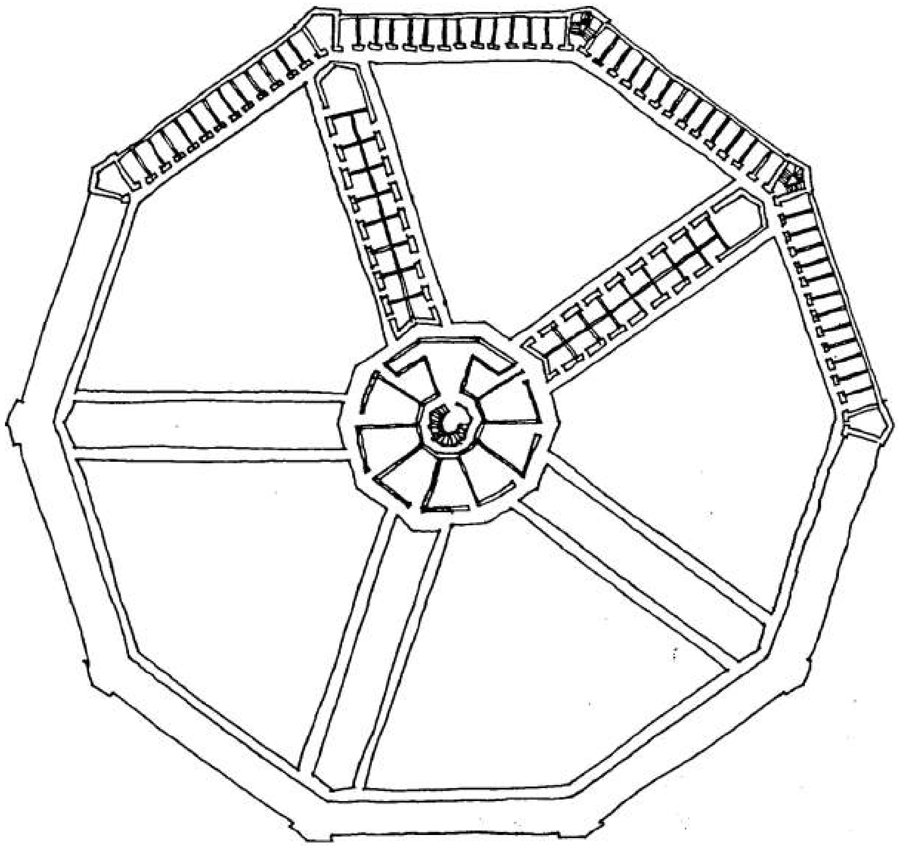


Figure 6.27 Design for a penitentiary for 600 prisoners (after Bevans 1819, prisons 3, plate 4).

a local differentiation of a continuous space structure with, characteristically, four ways out, which fully connect each street into its constitutive open rings. The same applies to a 'market place', which is the intersection of a pair of pairs of open rings, or a set of open rings. The 'square' is as natural in the unfolding of a 7-syntax as the 'bead' is in the 3-syntax.

4 Space and society

It may be objected that in giving this syntactic and largely abstract account of a theory of the formation of settlements and architectural complexes, we have ignored the customary lines of enquiry which seek to establish particular historical, economic or geographical 'causes' for particular patterns, or particular cases. It is not out of disrespect for this considerable body of work that we have, while

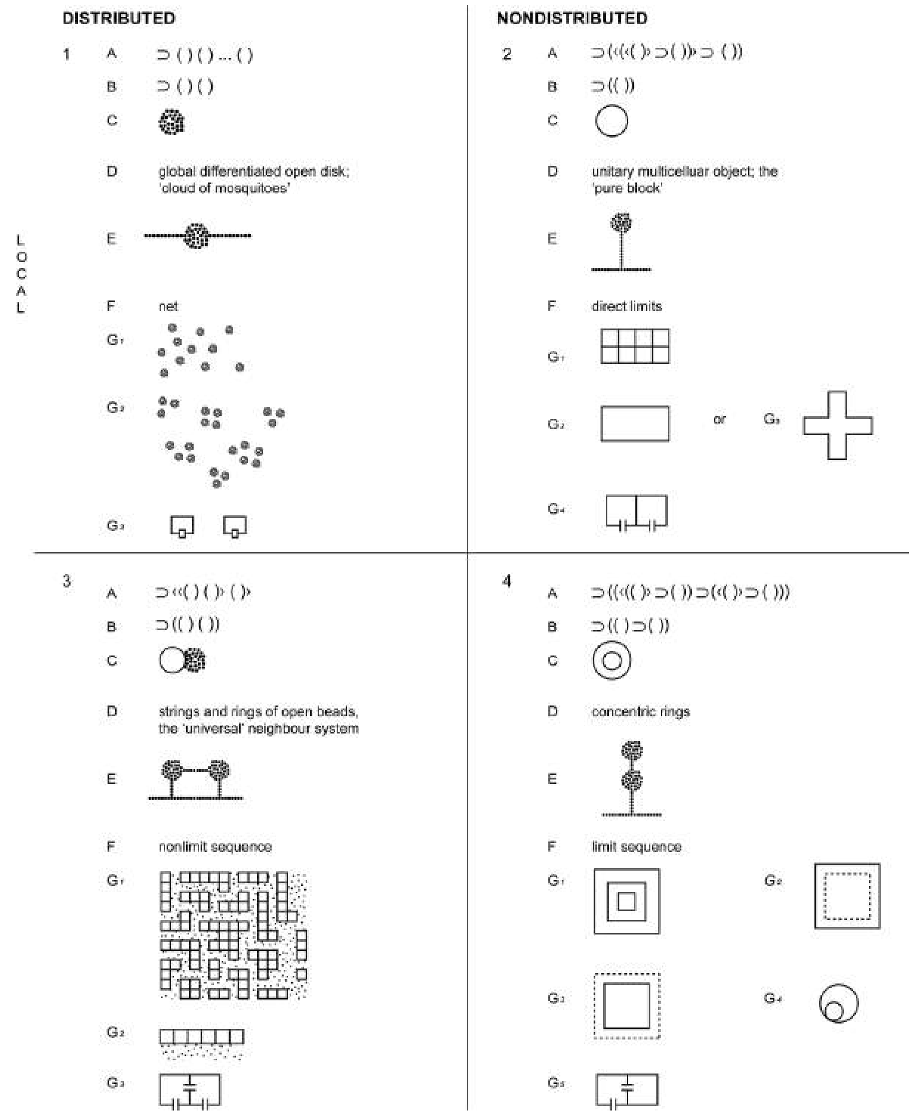


Figure 6.28 Morphological archetypes.

Key: A, generative formula; B, simplified formula; C, elementary space object; D, global space description; E, elementary path object; F, global path description; G, typical global objects and variants.

Notes:
(a) In all the syntaxes, the final example is the simplest realisation of the syntax for aggregates (that is at least two) of closed discs. All are extant morphologies, with known examples.
(b) *1-syntax:* G_2 repeats the 1-syntax at two levels; the smaller, a local zone, say a dwelling cluster of a minimal lineage; and the larger, the 1-syntax of these local zones in a global zone. *2-syntax:* G2 and G3 represent an alternative form of development for a 2-syntactic object; increasing the size of a simple closed disc while retaining the structure of the 2-syntactic object. The two examples are technologically possible ways of carrying out this development. In G2 the closed disc is stretched 'pairwise'; in G3 it is stretched 'double pairwise'. *3-syntax:* E, the path morphology results from the dense

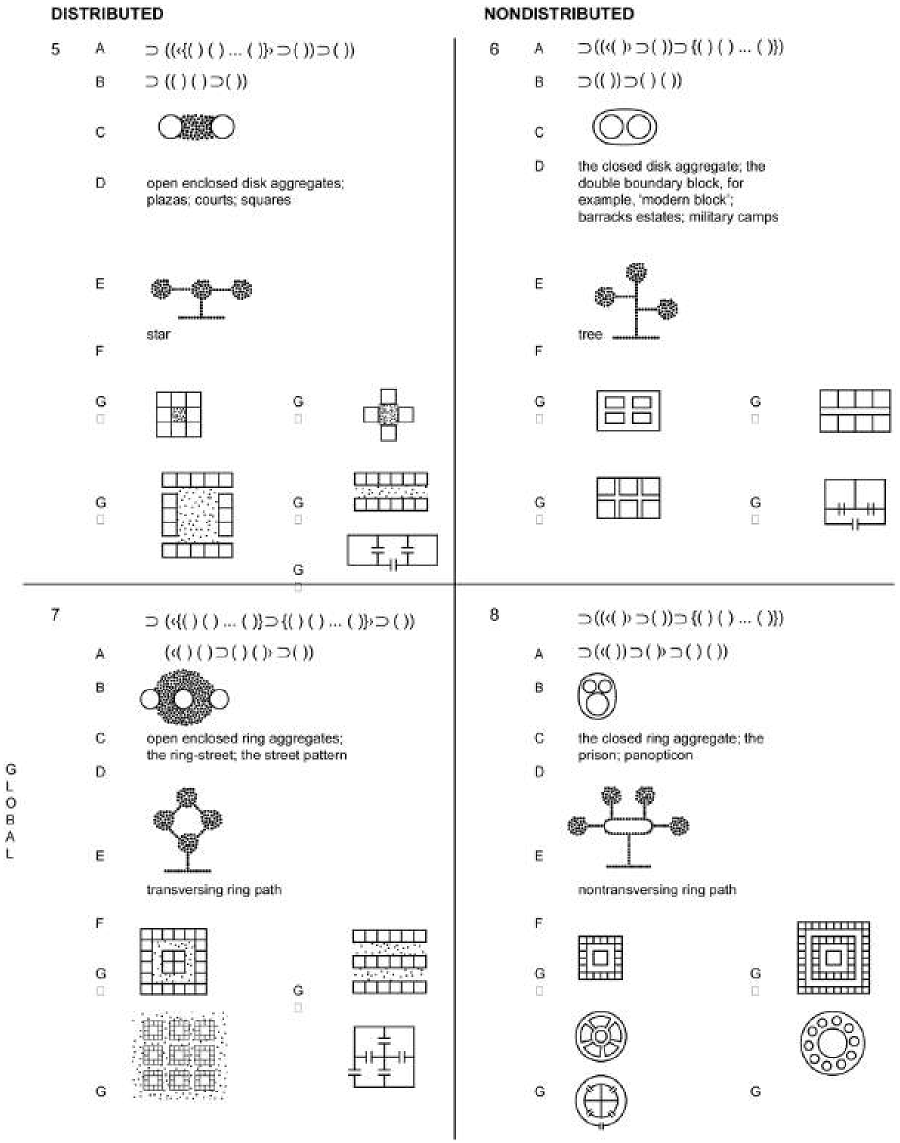


Figure 6.28 (continued).

aggregation of 'through permeable' objects. This implies that eventually there is another way back to the carrier path, hence the concept of a through permeable sequence, or nonlimit sequence. *4-syntax:* G, gives the simplest path realisation of the concept of a 'to-permeable' sequence, or limit sequence. *5-syntax:* E, the path morphology is a result of the aggregate controlling permeability to the open-enclosed disc, and thus to the aggregate of closed discs giving the 'star' form, and the 'inward-looking' aggregate. *6-syntax:* E, the 'tree' morphology interprets the concept of 'sequence' for a nontraversing path system. *7-syntax:* G_4, the aggregate resembles the 3-syntactic 'nonlimit sequence' but there is a difference. In 7, there is one space that is traversed twice in traversing the whole system, making a complete ring independent of the carrier path. *8-syntax:* G_4, this is drawn without its aggregates of small enclosed discs, because of the complexity of the form.
(c) Path forms here are related by formula analogy to the space forms. A parallel theory of path morphology is in preparation which is somewhat richer.

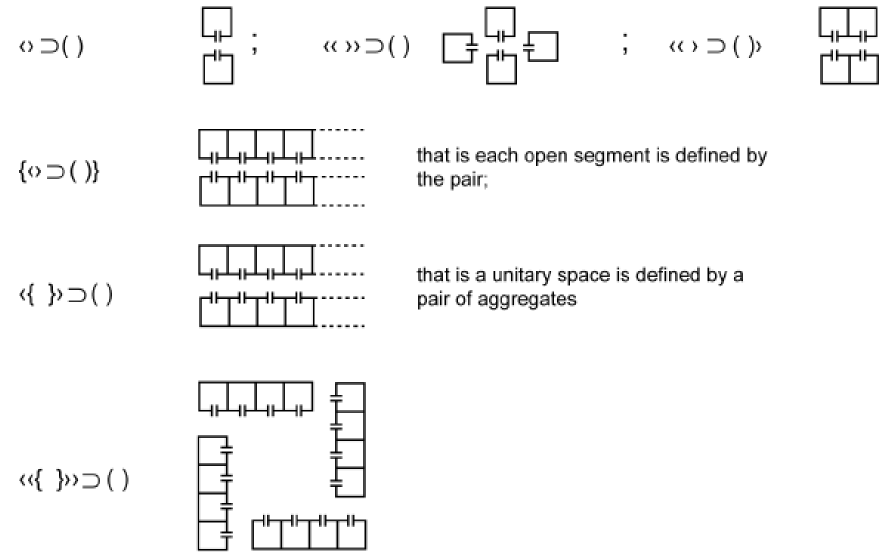


Figure 6.29 Examples of variants in the 5-syntax.

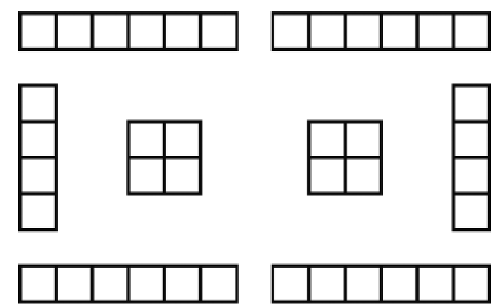


Figure 6.30 7-syntactic object with a pair of intersecting open rings.

assiduously analysing the existing record of real examples, temporarily adopted a resolutely abstract point of view. It is because of the widely acknowledged failure of descriptive and analytic work to reveal significant associations and relationships. We therefore took the view that until we knew what a settlement pattern was, and in what their essential similarities and differences consisted, it was futile to pursue causal explanations. We took the advice of Hermann Weyl:

> The experience of science accumulated in her own history has led to the recognition that evolution is far from being the basic principle of world understanding; it is the end rather than the beginning of an analysis of nature. Explanation of a phenomenon is not to be sought in its origin but in its

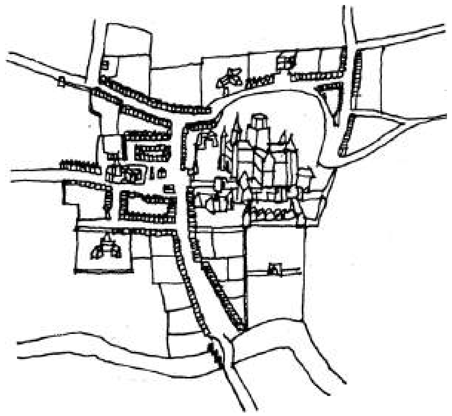


Figure 6.31 Early seventeenth-century Peterborough (after Speed 1974, plate XXIX, Figure 105).

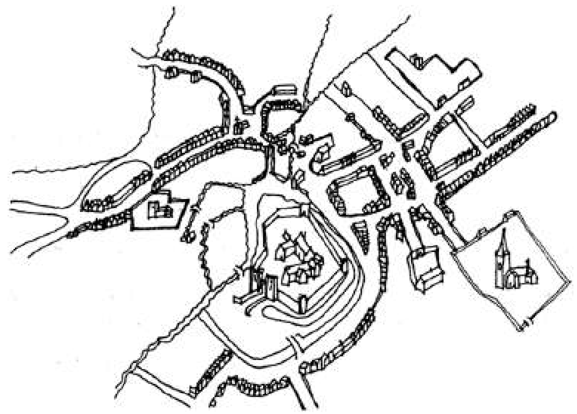


Figure 6.32 Early seventeenth-century Hertford (after Speed 1974, plate XXXI, Figure 110).

> immanent law. Knowledge of the laws and of the inner constitution of things must be far advanced before one may hope to understand or hypothetically to reconstruct their genesis.[33]

This is in our view clearly true in the study of the social production of architectural and settlement form. It is clear that in order to understand any particular example one needs to know two things: a particular set of historical determinants and contingencies; and the 'immanent laws' of the constitution of spatial structure itself. Our hope is that with at least a theory of the 'immanent laws' the empirical study of built form may be reinvigorated and perhaps to some extent more aptly directed. This anyway was the limit of our hopes when embarking on the development of the model by the careful analysis of the record.

However, as the model became better and better defined, it became clear that it might also be possible to associate with the space syntax model a social theory of the production and use of space patterns, by using similar concepts and methods to describe society as we had used to describe space. Almost as soon as this was attempted, a clear, if somewhat complex, 'inverse law' began to suggest itself as the relation between social and spatial structure. Furthermore, the theory as we developed it appeared to make sense and relate a significant proportion of the conclusions of other researchers, including architectural researchers, anthropologists, and even an economic anthropologist. Although it is in no sense a tested theory yet, having been developed and applied only retrospectively to the evidence collected by others, we are satisfied that the match between the formal structure of the theory and the distribution of evidence as currently known

is sufficiently suggestive and exact to permit our giving an account of it at this premature stage. Unfortunately, in the time so far available, it has not been possible to give a thorough review and reference to the work of others on which we draw. This serious deficiency will of course be corrected in a later paper.

This theory does not have a 'causal' form. It does not argue that particular forms of society 'cause' particular kinds of space patterns. Space is not a result of society as much as one of the means by which the social is constituted and made real. The theory is more in the form of *pattern similarities or relationships between spatial and social syntax-social* syntax being the patterns of encounters and relationships that hold among the members of a society. Nor is it correct to seek in such patterns and comparisons only a spatial *reflection* of social form. In certain cases, space is a reflection, but more usually it is an *alleviation* of, or *means* to, or even a *substitute* for social organisation. *Space is not a reflection of society, but a set of strategies in relation to social form, as often as not offering an alternative basis for encounters, other than those dictated by the social structure.* The streets of the city, for example, do not always reflect the social structure; they can be the means by which the pattern of social differences is forgotten and the inhomogeneous is assembled. They constitute the profane mixing of categories, which in the social structure are separate and insulated from each other.

In extending the concept of 'syntax' to social relationships and encounters, it must, of course, be stressed that no such syntactic theory yet exists on a level of exactness, comparable to the syntax theory of space. On the other hand, the theoretical ideas and descriptive work of certain anthropologists and sociologists is certainly responsive to such an interpretation. At a broad level, we shall try to show that both the general shape of the syntax model as a whole, and the patterns implied by particular syntaxes, provide useful ways of talking about social relationships. Before embarking on this, however, it is necessary to clarify exactly what the general shape of the syntax model is. For the following discussion, the reader is referred back to Figure 6.29.

The two columns divide the syntaxes into first, the *distributed*, or *glued together*, and second, the *nondistributed*, or *bound together*. In *distributed* syntaxes whatever space structure exists there is equally constituted by each primary cell. These are given odd numbers, and the higher the number, the more the syntax requires a global spatial rule (dominating the local placings) for its realisation. In *nondistributed* syntaxes the space structure is the result of an increasingly complex system of boundaries or spaces surrounding cells. These are given even numbers, and in the most local case the boundary simply encloses a continuous space, whereas in the high-numbered more global cases, boundary structures dominate the primary cells. In distributed syntaxes the integrating entity is always *inside* the collection of primary cells, and in some sense contained by them; in nondistributed syntaxes the integrating entity is *outside* the primary cells, and in some sense contains them.

Translated into social terms, a social order based on the *division of labour* (such as existed before the industrial revolution separated the skilled worker from his tools), in which each individual participates in society primarily through his functional interdependence with individuals possessing other specialities, is a concept which is both *local* and *distributed*. The global order arises out of a local ordering, that is, a pattern in which a particular individual repeatedly does a particular job. By contrast, the form of society that usually precedes this is that based on an elaborate and ritualised kinship system, usually involving a system of extensive naming of segments that are essentially similar (that is they all carry out the same functions). This is a *globally* ordered *distributed* system. The system still depends on being continuously recreated by the action of individuals, but these are controlled by a previously established global model. The global model in no way *arises* from the local actions; rather the latter conforms to the former. The former kind of society we could call 'urban', and the latter 'tribal'.

These two forms of society were called 'organically solid' (division of labour) and 'mechanically solid' (kinship) by the sociologist Émile Durkheim.[34] The latter form depends on a model which is not only global, but also symbolic; whereas the former depends on a model which is both local and instrumental – that is, it depends on real work, as opposed to symbolic work. Also, the former depends on real differences between people, whereas the latter depends on differences introduced by naming. For this reason, kin-based societies are sometimes known as 'segmental' because they are made up of large numbers of virtually identical segments.

These two polar types of social pattern appear to be *inversely related* to their corresponding spatial models on the local-global dimension, but *directly related* on the distributed-nondistributed dimension. According to the American anthropologist Elman Service, the ritualisation of kinship systems as a basis for social solidarity increases as the basic settlement units become more dispersed,[35] that is, at the 1-syntactic level. The less space physically integrates the society, the more integration depends on a global social model of a nonphysical (that is symbolic) kind.[36]

Exactly the opposite is true for the 'division of labour' form of traditional urban social solidarity. The theatre in which the division of labour (the local-to-global distributed social model) develops is the physically integrated space of the urban street pattern (that is, the global distributed model of space based on the 7-syntax). At the same time, the spatial model is physical, as opposed to symbolic, corresponding to the transition from kinship to the division of labour itself. From this follows the frequently observed association between the transition from kin-based to space-based society, and the transition from segmental work patterns to the division of labour.

In each of these polar cases, space plays an *inverse* role to the social structure. It is almost as though, at the 1-syntactic level, space provides a means of escape from the homogeneous social pattern[37], whereas at the 7-syntactic level, it integrates what has become socially differentiated. In both cases space alleviates social structure rather than reflecting it, yet is nonetheless systematically related to it.

The 3- and 5-syntaxes, on the other hand, have more closely parallel relationships between social and spatial structure, though in different ways. The 3-syntax, in which the global spatial pattern arises from local and distributed actions, corresponds to a social pattern which it formally resembles. This is the form of small-scale, distributed societies characterised by Bailey as 'multiplex'.[38] In brief, these are small spatially integrated societies in which each person is likely to know and encounter each other person for several different reasons. For example, the same person may be encountered as someone who serves you in a shop, whom you meet in a public house, who is your cousin, and who repairs your car in his spare time. Such encounter patterns are 'multiplex' in contrast to the type of encounter pattern generated by modern 'estates', where most encounters are specialised and not reduplicated in other areas of life. The theory is that 'multiplex' encounter patterns invoke the 'whole person' in continuous confrontations and, as a result, 'reputation' becomes of vital importance – and much social life is concerned with the negotiation and renegotiation of reputations. An unfortunate encounter in one domain of life will reverberate through all others and affect the whole 'reputation' of the person.

Like kinship patterns, the pattern of 'reputations' is still a symbolic reality, but it is no longer determined by some preestablished global model. It is continuously constructed by the negotiation of individuals. The global 'reputation pattern' at any particular time *is* constitutive of the social for that society, but it has been arrived at and is about to be changed by a collection of distributed local actions. This fluid, but strong, pattern is the same both for social and spatial patterns, except that the physical integration of space has accompanied the descent from a global to a local symbolic order. This is fully consistent with the overall pattern of development from kinship to urban societies.

A settlement in 5-syntactic form-buildings grouped around a central space is usually taken to be a case where the spatial form 'reflects' the social form in some sense. In our terms, a correspondence would be held to exist between a global and distributed spatial form, and a global and distributed social form. Research that exists on these societies does not support such a conclusion. Lévi-Strauss, for example, hints that the spatial form can be almost a disguise for the real social structure.[39] It represents a unity and simplicity of organisation, which is not possessed by the social structure itself. In such cases, it might be argued that the settlement pattern *represents* the society, but does not reflect its structure. Again this is consistent with the basic theoretical shift. Space increasingly provides an alternative basis in everyday life for a social structure whose complexities cannot be sustained in everyday practical life. Such settlements appear to be characterised by strong spatial and temporal categorisations of sacred and profane, and a tendency for these categorisations to play an important role in everyday and ritualistic life.

The nondistributed syntaxes exhibit virtually the opposite movement. The simplest nondistributed spatial gesture, the creation of a closed cell by means of a boundary (the 2-syntax), *established a domain of nondistributed spatial control* *within which the social takes precedence over the spatial*. If this is thought of at the level of the individual and his boundary (for example, a room), then we have something close to 'territorial' behaviour. Within the boundary, a social, local and nondistributed – but in all events strong – social model prevails. This is as true of the individual with his guest in the room, as of the open-plan factory, the open-plan school, the church and the football field.

At the other extreme the 8-syntax; although a global, nondistributed boundary system maximally controls the primary space, the spatial form totally dominates the social form and acts as a substitute for it. A prison is not simply about spatial control. It is about the elimination of social structure by the segregation of individuals (three to a room is a defect in reality, not in the theory!). A prison substitutes a globally defined, locally dominant, nondistributed spatial order for lateral social structure. A prison is a large, but essentially simple, social organisation. Its only social form is simple hierarchy (officially, that is – but this is why all films about prison life are centred around the prisoners' self-generated, informal social organisation). Otherwise, it has become homogeneous and segmented with the individual in his cell as the ultimate segment.

The 6-syntactic urban landscape of today is a milder form, but along the same lines: increasingly a spatial order is substituted for a social order, and this social order becomes a set of homogeneous, separate segments called nuclear families, with very strong sanction against the extension of social complexity even in the direction of a slightly extended family. Every activity has its own spatial boundary, and, correspondingly, social encounters are highly specific, rarely multiplex. Both space and encounter patterns are dominated by nondistributed agencies known as bureaucracies. Social life, other than that represented in the set of bounded locations permitted by the space pattern, is deterred both by the social and by the spatial pattern. The problem with this syntax is that it *does* reflect society. Indeed, its coerciveness is in no small part owed to the similarity of social and spatial syntax, which constantly reinforce each other to the point of appearing natural.

The remaining nondistributed syntax, the 4-syntax, is the other primary form of spatial order in the modern English landscape: the suburb. This is a local, nondistributed ordering based on a primary cell with a double boundary within which symbolic objects are placed (wishing wells, sundials, flowers) which express individual participation in a symbolic social order.

Syntaxes 3 and 6, and 4 and 5, have an interesting set of mirror relations. If 3 and 6 reflect social order, 4 and 5 appear to *misrepresent* it. Syntax 5 represents a simple global model of society, simpler than the social structure and perhaps more mythical than real. Conversely, syntax 4 represents an act of individual separation from society, which is again mythical. The spatial gesture of the suburb, with all its powerful sanctions to conform to an established pattern, sets up a myth of individual freedom and difference around an act of conformity and consensus.

These arguments may be summarised in the following general propositions:

(a) at the lowest syntactic level, distributed *space is a means of escape* from the social;
(b) and at this level, nondistributed space constitutes a minimal domain within which the *social prevails over the spatial*;
(c) in general, both for distributed and for nondistributed syntaxes, *space increasingly becomes an alternative basis for the social*; but
(d) if distributed, the higher-number syntaxes put together in space what is socially differentiated;
(e) and if nondistributed, they separate what is socially the same, substituting a spatial regime for a social; and in general
(f) distributed forms constitute an alternative socialness in spite of social inhomogeneity (for example, the relation between the integration of urban space and the division of labour);
(g) and nondistributed forms substitute spatial control for social complexity and inhomogeneity; at the broadest level
(h) low-numbered distributed syntaxes are associated with socialities which are small and homogeneous; low-numbered nondistributed syntaxes with social organisation that is small and internally complex; high-numbered distributed syntaxes with societies that are large and complex (that is inhomogeneous); and high-numbered nondistributed syntaxes with social forms that are large and simple, that is, both segmental and hierarchical, but lacking complex relations among members.

5 The analysis of real domains

These broad relationships serve as a useful backcloth to the analysis of real socio-spatial patterns, but they are only a stepping-stone to the methodology we need to deal with spatial *processes* and *transformations*. Perhaps paradoxically it is at this point that the problem of *description* must be revived in relation to processes of real domain constructions.

A real domain is a $\bar{\sqsupset}$ relation, or a set of such relations, to a carrier space. A subdomain is a domain whose carrier space is itself a domain. A real domain may be, or become, a stronger or weaker realisation of a certain syntactic type. For example, if the subdomains constructing a street pattern are progressively replaced by blocks of subdomains controlled by a single entrance, then the domain relations of the street (or 7-syntax) are progressively removed, and the domain becomes more and more weakly a 7-domain. This corresponds to an intuitive effect that is usually explained in terms of 'scale' but which, as with many other 'scale' effects, is quite naturally explained as a syntactic effect.

Any domain at any scale (from a simple house to a settlement pattern) is constructed by a process which articulates two kinds of syntactic structure: the *transformation* structure, which gives the nature of units; and the *combinational* structure, which relates each unit to other units. A domain may have k such *interfaces*, that is, $k+1$ levels of syntactic organisation. These interfaces, rather than the levels taken 'independently', appear to be the key to the transformational analysis of real domains that are not characterised by a simple syntactic process. In most cases one interface in a process will be more important than others, and will be called the *dominant interface*.

To begin with, we may use the notions of description and description retrieval to distinguish *natural* and *unnatural* domain-processes. An unnatural process is one in which a description retrieval has intervened to introduce more global order into the process. For example, a 3-syntactic aggregate will sooner or later generate an asymmetrical open ring, whose description may be retrieved and introduced as a global order for the next stage of growth. Exactly such a process would constitute a minimal town in which the 'market place' was constituted by the intersection of the pair of open rings. A natural process is one in which description retrieval does not intervene to increase the level of order, although there exist natural processes which also produce more syntactic order as the aggregate grows large.[40]

In general, as a domain grows it poses problems of description retrieval, which are essentially problems of control, and which normally require more global thinking (that is, conscious design) for their solution. In fact, contrary to current romantic theories of the vernacular, conscious design intervenes in almost all aggregates above a certain small size. In particular, two kinds of description retrieval problems are critical: those concerning the *relations between social and spatial organisation*; and those concerning the *relations across the dominant interface*, that is, between *transformation* and *combination* structures of subdomain and domain.

An apparent general property of domain processes concerns all of these: *the larger a compact spatial aggregate becomes, the stronger must be the social structure which relates it to comparable aggregates across the carrier space*. The converse of this is a general proposition, argued by Sahlins,[41] that spatial fission occurs in the 'state of nature' to avoid the construction of an over-strong social structure.

To illustrate the proposition itself, we may refer to recent work by Bradfield.[42] Among the Tallensi, the compact aggregates are small, familial compounds which never grow above a certain size. In such a case, a relatively weak social structure is adequate at the combination level, consisting more of symbolic and ritualistic arrangements than explicit sanctions. The villages of the Mende, on the other hand, where the compact aggregate is much larger, have much stronger secret societies (which Bradfield suspects may have to do with the emergence of social classes) which operate largely at the level of relations between villages. When considering towns, this development reaches a new level. The exigencies of relationships between settlements are such as to transform the within-settlement social structure into the embryonic form of a class structure.

It may be speculated, on the basis of this proposition, that tribal and urban societies are not, after all, stages on the same evolutionary pathway, but divergent socio-spatial processes which occur from the beginning of agriculture. Tribal forms are essentially based on distributed, noncompact, noncontiguous syntaxes,[43] with space in a largely symbolic role, and social structures constructed on a symbolic basis without strong sanctions. The nondistributed version would be tribal conquest systems, in which an instrumental order of man-to-man relations predominate over an expressive order, and nondistributed settlements control a large landscape (see, for example, the Zulu kraal in Figure 6.26). Urban societies are essentially based on the primacy of man-to-nature relations, the division of labour, the compacting of space, and consequent increase in the strength of a sanctions-based social order.[44] The earliest distributed versions of this are in ancient Mesopotamia and the nondistributed versions are in pre-Columbian America, for example, the Aztecs.[45] Against this background, it would perhaps be useful to reconsider feudalism as a socio-spatial form, with particular attention to the dominant interface.

These are preliminary considerations, however. The only justification for including them in the paper is to show the potential usefulness of both a formal approach and a socio-spatial framework to the analysis of social as well as spatial forms.

Acknowledgement

This is a revised version of a paper submitted to the Science Research Council reporting the initial stage of a research programme supported in full by SRC Research Grant GR/A/1769.

Notes

1. E. Leach, 'Discussion: Does space syntax really "constitute the social"?', in *Social organisation and settlement: Contributions from anthropology, archaeology and geography*, edited by D. Green, C. Haselgrove and M. Spriggs (Oxford: British Archaeological Reports, 1978), pp. 385–401; B. Hillier, A. Leaman, P. Stansall and M. Bedford, 'Reply to Professor Leach', in ibid., pp. 403–405.
2. B. Hillier and J. Hanson, *The social logic of space* (Cambridge: Cambridge University Press, 1984).
3. B. Hillier and A. Leaman, 'The architecture of architecture: Foundations of a mathematical theory of artificial space', in *Models and systems in architecture and building*, edited by D Hawkes (Hornby: Construction Press, 1975), pp. 5–28.
4. A. N. Whitehead, 'Mathematics and the good', in *The interpretations of science: Selected essays*, edited by A. H. Johnson (Indianapolis: Bobbs-Merrill, 1961), pp. 187–203.
5. D. Michie, *On machine intelligence* (Edinburgh: Edinburgh University Press, 1974), p. 117.
6. D. Michie, *On machine intelligence*, p. 141.
7. Perhaps space-time would be more accurate, in which case the expression 'carrier space-time' would follow.
8. I. Hacking, *Logic of statistical inference* (Cambridge: Cambridge University Press, 1965).
9. Editors' note: an Erratum was added to a printout of the article in the form of a pasted, typed text, stating 'Mathematical symbols strip the morphic unit of all its particular properties, and leave only the most abstract and universal properties – of being a member of a set, of existing and so on.' We are grateful to Paul Stansall for providing this text in February 2023.
10. B. Bernstein, *Class, codes & control* (St Albans: Paladin, 1972).
11. Editors' note: we have renumbered – and where necessary, formulated new numbers and captions – for all the figures in the text. Many of the figures in the original text were missing their numbers and/or captions. The figures have been recompiled and renumbered accordingly.

12. M. A. Arbib, 'Self-reproducing automata: Some implications for theoretical biology', in *Towards a theoretical biology*, edited by C. H. Waddington (Edinburgh: Edinburgh University Press, 1969), pp. 204–226.
13. H. Weyl, *Philosophy of mathematics and natural science* (New York: Atheneum, 1963), p. 147.
14. At this point in the text subscripts are discontinued.
15. The term 'operation' will be used when relations and 'numbers' of objects are specified in a general way. When specific objects are added, such that the formula expresses a definite morphology, the term 'rule structure' or simply 'rule' is appropriate.
16. This is, if the historical evidence and our syntactic interpretation of it are to be believed, the origin of the concept of a complex building – the analysis of complex building being the analysis of the space and path structures that can be defined on such a structure.
17. In the following discussion the reader is advised to refer to Figure 6.28.
18. R. Thom, *Structural stability and morphogenesis: An outline of a general theory of models*, translated by D. H. Fowler (Reading, MA: Benjamin, 1975) (1st Eng. edn).
19. M. Fortes, *The dynamics of clanship among the Tallensi: Being the first part of an analysis of the social structure of a Trans-Volta tribe by Meyer Fortes* (London: Oxford University Press, 1945), p. 155.
20. Image based on L. Prussin, *Architecture in Northern Ghana* (Berkeley: University of California Press, 1969), pp. 56, 59.
21. For other examples of this, or similar patterns, see for example E. Z. Vogt, 'Some aspects of Zinacantan settlement patterns and ceremonial organization', in *Settlement archaeology*, edited by K. Chang (California: National Press Books, 1968), pp. 154–173 on Zincantan in Mesoamerica; and also the observations of M. Sahlins, *Stone Age economics* (London: Tavistock Publications, 1974).
22. F. H. H. Roberts, *Ruins at Kiatuthlanna Eastern Arizona*, Bureau of American Ethnology Bulletin 100 (Washington, DC: Smithsonian Institution, 1931).
23. J. Mellart, *Çatal Hüyük: A Neolithic town in Anatolia* (London: Thames and Hudson, 1967).
24. Images after OSC, 1888–1893, first edition, Ordnance Survey Collection, British Museum, London.
25. Images after OSC, 1888–1893.
26. Image after T. E. Peet, J. D. S. Pendlebury, J. Černý, H. Frankfort and L. Woolley, *The city of Akhenaten*, Memoirs of the Egypt Exploration Society, 38, 40, 44 (London: Egypt Exploration Society, 1923).
27. OSC, 1955, TQ3184, 25-inch scale map of Greater London, Geography Library, Senate House, University of London, London.
28. Image after HMSO, Conwy Castle and Town Walls (London: HMSO, 1957).
29. E. K. Reed, 'Types of village-plan layouts in the southwest', in *Settlement patterns in the New World*, edited by G. R. Willey (New York: Wenner-Gren Foundation, 1956), pp. 15–16.
30. S. A. Stubbs, *Bird's-eye view of the Pueblos* (Norman: University of Oklahoma Press, 1950).
31. J. Bevans, 'Design for a penitentiary or gaol for 600 prisoners', in *Reports from Select Committees on Prisons and Gaols 1819–1836 Parliamentary Papers* (London: House of Commons, 1819).
32. Figures after illustrations by Speed in P. Lavedan and J. Hugueney, eds, *L'urbanisme au moyen age* (Geneva: Bibliotheque de la Societe Francaise d'Archéologie, 1974), plates XXIX, XXXI; figures 105, 110, respectively.
33. H. Weyl, *Philosophy of mathematics and natural science* (New York: Atheneum, 1963), p. 286.
34. É. Durkheim, *The division of labour in society*, translated by G. Simpson (New York: The Free Press, 1933) (1964 Eng. edn).
35. E. R. Service, *Primitive social organization: An evolutionary perspective* (New York: Random House, 1971).
36. This is reflected in two dominant space-codes in our society. In general, the middle class have aspatial networks, and relationships are constituted by patterns of ceremonial occasions, in particular, inviting people to dinner. The traditional working class, by contrast, have strongly spatial (local) networks with much freer general access to dwellings, but a tacit prohibition at mealtime, which is a private occasion.
37. M. Sahlins, *Stone Age economics* (London: Tavistock Publications, 1974).
38. F. G. Bailey, *Gifts and poison* (Oxford: Blackwell, 1972).
39. C. Lévi-Strauss, 'Do dual organizations exist?', in *Structural anthropology* (Harmondsworth: Penguin, 1972), pp. 132–156.
40. These processes are the subject of a forthcoming programme of computer experiments.
41. Sahlins, *Stone Age economics*.
42. R. M. Bradfield, *A natural history of associations: A study in the meaning of community*, vol. 1 (London: Duckworth, 1973), Chapters 1–5.
43. Current theory suggests that there may exist a mirror set of *negative syntaxes* which are, in essence, syntax theory interpreted for point arrangements, and which may be called 'negative' by analogy with negative numbers, with the first syntax as the 'zero' of positive and negative syntax.
44. G. Park, *The idea of social structure* (New York: Anchor Books, 1974).
45. An as yet unpublished study of this theme has been undertaken by Ross Donaldson at the School of Environmental Studies, University College London.

7 Space syntax: A different urban perspective (1983)

The impact of space syntax: A pivotal point for design. Introduction to 'Space syntax: A different urban perspective'

Ricky Burdett

The notion of 'impact' for academic research did not feature as aggressively as it does today when this article was published in 1983, 40 years ago. Yet, for the rarified theory of space syntax, it was its first and one of its most impactful outings. The article extended over 16 heavily illustrated pages – with legible maps, diagrams and photographs – in *The Architects' Journal*, one of the most popular weekly magazines read by architects and designers in the UK and abroad. It marked a pivotal turning point where pure research touched the real world.

More than a decade later, one of the world's most respected architects, Norman Foster, claimed that the experience with Bill Hillier and his team at UCL had been instrumental in shaping his vision for numerous urban projects. In his view, he declared in 1997, space syntax allowed experiments in the interaction between the world of 'analysis, observation, reason and research' and the world of 'passion, feeling, intuition, imprecision and the hunch'.[1] Quite a claim for a relatively esoteric analytic theory developed by a Cambridge-educated English Literature graduate who 25 years earlier had been leading the Intelligence Unit at the Royal Institute of British Architects (RIBA) and was 'the nearest thing we have to a longhaired intellectual'.[2]

From his early days at the RIBA and during his leadership of the Unit for Architectural Studies at the Bartlett School of Architecture, Bill Hillier and his colleagues consistently sought to link theory to practice, albeit in rather academic ways. This approach informed many of the early research papers, including the 1976 article which presented space syntax as a theory of the fundamental topological principles that underpin the social production of built space, but it remained relatively abstract. Nonetheless, in the early 1980s the methodology was used in some design and legal settings in London to evaluate whether architectural proposals were integrated into their wider urban contexts – a key concern amongst community groups and planners of the period.

Space syntax was employed in a Public Inquiry (a UK government legal process) as 'scientific evidence' in support of a community-led project in East London's Limehouse and a critique of more commercial schemes. The full analysis was featured in the *Architects Journal* article as an indication of how space syntax could be used to provide an objective assessment of the spatial and social potential of architectural concepts (Norman Foster's 'intuition') to test its impact on an urban area. As an urbanist with experience in the field, I know of no other research methodology that has been employed in this manner over the last decades.

Importantly, the theory-led, practice-oriented approach shaped the teaching ethos of the master's programme at the Bartlett (the slightly misnamed MSc in Advanced *Architectural* Studies), which attracted an international range of emerging and mature design practitioners stimulated by the intellectual climate of the research and teaching unit. Where else could you learn about Karl Popper's theory of falsifiability and Claude Lévi-Strauss's structural anthropology and apply it to the discourse and practice of contemporary architecture? It is in this environment that the intersections between theory and practice were nurtured.

With its outsized computers and reams of pre-punched paper printouts scattered along the corridor at the opposite end of the building occupied by trendy architecture students in an open-plan studio, the Unit became a physical and virtual meeting point of individuals and ideas that linked the social to the spatial. I count myself as one of the privileged global fraternity which benefitted from the experience which shaped my own career. The early mapping and computational techniques developed by Bill Hillier and his researchers were evidence of a methodology 'in-the-making', in an embryonic stage. There was a buzz but no certainty on how this end of the corridor (research) would influence the other (architecture). Or, vice versa.

The publication of 'Space syntax: A different urban perspective' in the *Architects' Journal* (AJ) made a pivotal difference. As a published article, it punched way above its weight. Not because it set out a new theory or methodological innovation, but simply because it was conceived, written and designed to communicate. As the most junior of the authors, I remember the time, care and attention given by the production team at the AJ to transform space syntax's rudimentary mapping techniques (rough maps, hand-drawn axial lines, wobbly convex space, messy depth diagrams) into elegant and intelligible overlays that were effective in rendering abstract ideas concrete. Every phrase that Bill Hillier took for granted ('social modality', 'genotype', 'axiality', 'co-presence', 'integration core') was questioned, examined, simplified and translated into a language that every non-academic reader could understand.

To have real impact on the design professions, the editors foregrounded the comparative analysis between a much-loved (by the English), seemingly unplanned French medieval town of Apt in Provence with a recent award-winning housing project in a tough part of North London – the Marquess Road estate. It was a smart choice at many levels. It connected with the language and fashionable terminology around concepts of defensible space and community architecture. It went to the heart of Foster's invocation for 'passion, feeling, intuition, imprecision and the hunch', which is how architects think and sketch. At a visual level, the contemporary design of the estate adopted a soft, vernacular style with brick facades and front porches reminiscent of its continental predecessor. Its layout was based on courtyards and alleyways – not elevated concrete walkways and access balconies – also inspired by traditional urban forms. The architects Darbourne and Darke were sensitive housing designers with a strong track record of socially responsive projects that rejected the formulaic geometry of so many Modern Movement projects influenced by the Charter of Athens. Yet Marquess Road had failed as a social experiment. It was dark, empty and soulless, especially in its more central areas. Space syntax was there to explain why!

The central methodological argument presented in the article hinged on the important distinction between a 'local' and 'global' analytical perspective. This was new. A profoundly structuralist approach, it both applauded and critiqued the popular visual 'townscape' narrative so ably articulated by the seductive, picturesque sketches developed by Gordon Cullen, a highly influential architectural commentator of the 1960s and 1970s. While these 'townscape' images captured the three-dimensional visual qualities of an urban space as seen by the human eye, they failed to capture the overriding spatial order that determines the 'social logic' of the pedestrian movement dynamics of the wider urban system. Let us remember that neither Kevin Lynch nor Jane Jacobs had been able to capture the complexity of this metropolitan dimension of urbanity.

At a more pragmatic level, the article's key methodological innovation was the representation of urban systems through a relational system of axial lines and convex spaces, which respectively defined the global and local properties of a human habitat, independent of cultural, topographic or social context. Convex and line-based representations are translated into graphs, and graphs are drawn or 'justified' to visually capture the way in which individual spaces are more or less accessible from other spaces (in terms of changes of direction). This allowed the graphic representation of urban space as a system of relational differences. The mathematical modelling of the system produced an *integration core* which represents how *accessible* a space is from all other spaces in the wider area. Without repeating the article's conclusions, the syntactic analysis of the Marquess Road estate demonstrates that its integration core remains on the periphery, failing to link the interior to its wider neighbourhood. A similar critique is applied to three competing schemes for a major regeneration development in East London's Limehouse area, making a strong case in support of a community-backed scheme that invites people in rather than keeps them out.

While the intellectual framing for such a powerful and original methodological leap had been developed in earlier academic articles and would be fully laid out in the soon-to-be published *The social logic of space* (Hillier and Hanson 1984), the *Architects' Journal* text lays it all out on the table: a 'different urban perspective', as the subtitle exclaims. Hillier and his colleagues were to later develop the argument further, taking into account the discovery that the syntax of urban space plays a key role in determining the distribution of pedestrian movement and co-presence in everyday life.

In hindsight, it could be argued that if this article had not appeared in the professional press, Norman Foster and many others may never have come across space syntax and it would not have become the pervasive global 'design' tool that it is today. The success of the pioneering work of the Unit for Architectural Studies and its many participants was founded on its intellectual and methodological underpinnings established by Bill Hillier, but this one publication perhaps helped secure its engagement with the transformation of our built urban reality.

Space syntax

A different urban perspective

Bill Hillier, Julienne Hanson, John Peponis, John Hudson and Richard Burdett

Originally published: Hillier, B., Hanson, J., Peponis, J., Hudson, J. and Burdett, R. 1983. 'Space syntax: A different urban perspective'. *The Architects' Journal*, 178(48), 47–54, 59–63.

Architects may have given up the idea that 'community' can be created by design, but the question of whether architecture can have any social effects is still as open as ever. After several years of research funded by the SERC,[3] Bill Hillier, Julienne Hanson, John Peponis, John Hudson and Richard Burdett of the Unit for Architectural Studies at the Bartlett have a modest but nonetheless fundamental proposition to make in answer to that question. If true, it could have far-reaching implications for urban design and redevelopment.[4]

Demise of conscious design

Most architects today believe that something has gone badly wrong with the design of urban space. No matter how hard they try, they do not seem able to recreate the unforced, informal liveliness that once contributed so much to the quality of urban living.

The search for 'urbanity' has become a central theme in architecture, underlying a growing variety of contending styles and movements, just as it has become a major preoccupation of the increasing number of public inquiries into urban redevelopment. The neo-vernacular urban villages (Odhams Walk in Covent Garden, the Marquess Road estate in Islington), and the grander neo-historicist gestures of Aldo Rossi, Ricardo Bofill and Robert and Leon Krier, are as much about this as are the show-trial inquiries of Coin Street and Limehouse Basin on the Thames.

For many people, the problem is not architecture but the lack of it, and it is seen as an institutional problem. The architecture of urban space has been lost in the interregnum between architecture and planning. Piecemeal rules and regulations have taken over from conscious design. The main blame, according to those who support this argument, lies either with the planners' insistence on 'zoning' principles or with the traffic engineer and the impossibility of reconciling urban living with the motor car. To some extent this is undeniable. But it cannot be the whole story. On the one hand, we are surrounded by inherited urban environments which have adapted perfectly well to modern living, and on the other it seems exactly where architecture has made its most strenuous efforts to reinterpret urban space that the most notorious 'urban deserts' are said to be, from the 'streets-in-the-air' and 'urban rooms' of 'social Modernism' to the unoccupied village greens and alleys of toy-town vernacular. One reaction to failures in post-war urban design is the urban-village school of design. Despite good intentions, the result is essentially an urban desert. Pictured: Marquess Road estate, image from John Peponis's archive circa 1983 (Figure 7.1).

The problem, it must be admitted, is one of *knowledge* – *architectural* knowledge. There is a substantial gap in our knowledge of the social implications of strictly formal, hence architectural, decisions. There is no adequate description and explanation of why certain types of spatial patterning seem, inevitably, to lead to that curious feeling of a disembodied architecture, devoid of human contact and activity, any more than there is an understanding of why the urban space of the past so easily provided a setting for the life that nowadays seems so often to be missing.

The conceptual problem

We believe that this lack of knowledge – and probably the loss of responsibility as well – stems from a conceptual difficulty. Designers do not have concepts and techniques that allow them to describe and investigate the kinds of spatial order that are to be found in highly complex physical objects like towns and cities. It is because designers today do not properly understand their *spatial* logic that they cannot develop a proper understanding of their *social* consequences.

The problem is one that architecture shares with many other branches of knowledge: that of understanding patterns of spatial relationships. All architecture is about such patterns and it is very difficult even to describe them in a way that helps our understanding. In urban space a Gordon Cullen sketch can show us pattern properties that can be seen from a single point.[5] But urban space – indeed all architectural space – is something that must be understood from many points if we are to understand its social nature and consequences. Visual representations, such as Cullen's, give the designer too *local* a view, whereas we argue that the *global* properties of spatial patterns must be understood. In architecture this is, as we can now show, a strictly practical question, since the main finding of our research into urban space is that it is the global organisation of space that acts as

Figure 7.1 Marquess Road estate. Credit: John Peponis.

the means by which towns and urban areas may become powerful mechanisms to generate, sustain and control patterns of movement of people. Our research has shown that spatial organisation – over and above any effects due to the location of facilities and population density – has a crucial effect on the ways people move through an urban area, and therefore on the ways people become automatically aware of each other.

Social well-being

The awareness of others is not perhaps what sociologists have called 'community'; it is more like a latent or 'virtual' community which we have come to believe is important for its own sake, because it offers a sense of safety and belonging which may flower into a community. Whatever its sociological status, this virtual community is in fact the distinctively *architectural* contribution made to social well-being.

The fundamental problem of recent urban design is that, no matter how it is disguised, it follows an *enclave* philosophy which promotes the design of enclosed, inward-looking clusters of houses. This attitude is so deeply rooted that it is often taken for granted, especially since Oscar Newman's 'defensible space' principle – founded on the dubious territoriality theory – persuaded many that it had some foundation in scientific fact.

We are now sure that, however designed, no architectural philosophy of enclaves can solve the problem of recreating urbanity. All enclave architecture reflects an over-localised conception of design. How buildings are arranged around particular spaces is important, but this form of spatial arrangement can never reproduce urbanity. How a space fits into an area is a more important determinant. Urbanity and the virtual community are the products of the larger-scale organisation of space, that is, global design.

What is space syntax?

Space syntax is a method we have developed at the Bartlett Unit for Architectural Studies to describe and analyse patterns of architectural space – both at the building and urban level. The idea is that, with an objective and precise method of description, we can investigate how well environments work, rigorously relating social variables to architectural forms. We can thus simulate the performance of real and hypothetical schemes on the computer, so that it can be used as a suggestive and evaluative design tool.

Obviously the design tool will only be effective to the degree that we can establish general principles relating spatial form to social outcome, checking observations against computer simulations.

Spatial effects

Research using this method at the Bartlett involved the analysis and simulation of more than 100 towns, urban areas and design proposals, and the systematic observation of 15 examples. Clearly much remains to be investigated and explained. Nevertheless, our results show unequivocally that the spatial organisation of towns and urban areas affects patterns of movement and use according to well-defined principles, which relate to the *intelligibility* of space – that is, how easily inhabitants can distinguish between the larger pattern of space and the local parts; the *continuity of occupation* – that is, whether there are pockets of unused or underused space in an area; and the *predictability* of space – how well the potential pattern of encounter can be predicted from the spatial pattern (see the appendix at the end of this article). All these terms can be precisely defined and evaluated, but before doing this, we must explain the method of analysis.

Space syntax: Defining the method

First, we must define what we mean by spatial order, and here one comes up against a problem, namely that most ideas about spatial order are in general geometric, while many urban environments are not – they often appear to be in a kind of disorder. But geometric order and spatial order are not necessarily the same thing; their local and global properties may be quite different.

Deformed grid

A highly geometric scheme, such as Ricardo Bofill's 'new socialist village' in Algeria (Figure 7.2), looks like an intelligible order when seen from above in two dimensions, when it can be grasped as a whole. But it may not be intelligible when one moves about in it because parts are repeated. In contrast, traditional towns, such as Apt (Figure 7.3) in the south of France, are often characterised by an irregular non-orthogonal pattern, where the urban grid is 'deformed' so that it generates a rich and varied spatial quality. It is very difficult to see any geometrical order in such towns. Even towns like New York, which at first seems to be organised on a gridiron plan, often have the same spatial property as Apt, in they introduce a degree of urban grid deformity (Figure 7.4). It is the deformation of the grid that brings about local differentiation – which distinguishes the parts from each other – and provides the means by which inhabitants can assemble the parts into a coherent whole.

What do we mean by a deformed grid? First, compared to an orthogonal grid, the length of sightlines from particular spaces – their one-dimensional extension – is sometimes restricted and sometimes extended. This one-dimensional extension we call axiality. Second, the width of spaces – their two-dimensional extension – varies considerably. This we call convexity. In Apt, the buildings seem to be arranged in such a way as to create a continuous flow of open space with wider and narrower sections and shorter and longer perspectives. The 'streets' and wider spaces are always lined by entrances to buildings, leaving few areas of blank walls. Also, the layout offers a choice of routes from any point in the town to any other, with few cul-de-sacs. These rather vague descriptions can be precisely formulated graphically, as can be seen in Figures 7.5–7.6.

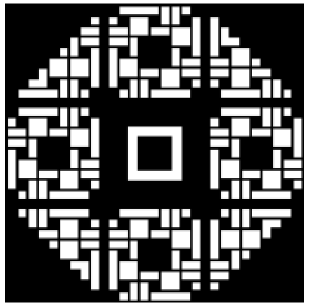


Figure 7.2 Ricardo Bofill's 'new socialist village' in Algeria.

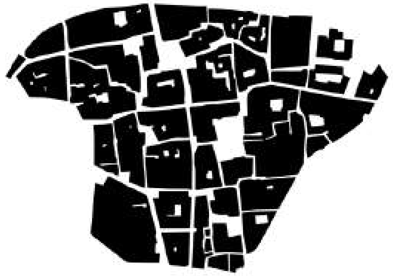


Figure 7.3 Apt, Vaucluse, in the south of France.

Figure 7.4 New York's gridiron plan.

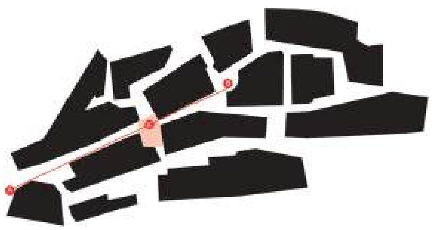


Figure 7.5 Any point in the open space of an urban area can be seen to extend linearly (or axially) in one dimension – the line from A to B – and also convexly in two dimensions (the toned area).

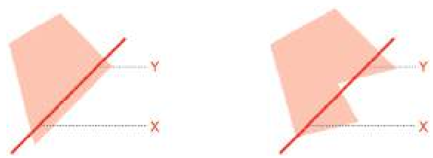


Figure 7.6 By definition, any convex segment will not contain concave parts. Any two points in a convex segment can be joined by a straight line which does not go outside the boundaries of the space.

Axiality and convexity will be shown to account for the way in which space – whatever its style or three-dimensional form – structures movement. This does not mean that the third dimension is not an important aspect of an architectural appreciation of urban space, but that the movement of people through an urban area is limited to two dimensions.

An axial map will consist of the fewest and longest straight lines that cover the entire surface of a town, taking account of how far you can see and how far you can walk. As a way of seeing and experiencing a town, an axial map offers the most 'globalising' perspective, since an axial line will extend as long as at least one point is visible and directly accessible from it. A convex map will be composed of the

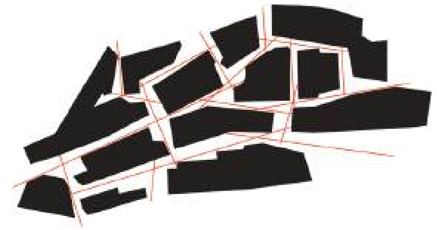


Figure 7.7 An axial map.

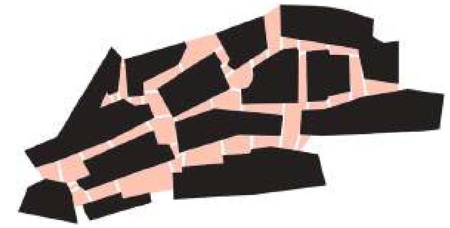


Figure 7.8 A convex map of the same settlement.

largest and fattest convex spaces that cover the entire area being analysed. A convex space is the most 'localised' because its boundary is defined by every point that is directly accessible to every other point in that convex space. (See Figures 7.7–7.8.)

Organisation of space

The idea that every point in the system has both a one- and a two-dimensional aspect is different from the idea of streets and squares, where spaces are expected to be either one- or two-dimensional. Seeing every point in both ways means that every point has both a local and a global dimension. It is how the two come together that distinguishes the different morphological features of different types of town and urban areas. Both the convex and axial organisation of space constitute an interface between the open space and the buildings. In Apt it is noticeable that nearly every convex space – and hence every axial line – is directly accessible to at least one building entrance. This in fact is a normal property of many traditional urban spaces, and the opposite of most recent neo-vernacular schemes, as seen at Marquess Road estate.

Urban comparison

At this stage it is useful to introduce a key distinction between the kinds of people who are affected by the physical arrangement of space: the inhabitants (who live beside or near a particular group of spaces), and the strangers (who do not belong to a particular set of spaces but pass through *en route* to another area).

Inhabitants and strangers

Although the presence of strangers is generally accepted as being crucial in creating an awareness of others and liveliness in urban areas, it also plays an important role in policing space. Unlike Oscar Newman's 'defensible space' theory, which emphasises inhabitants policing space and excluding strangers, our research has led us to conclude that strangers police space and inhabitants police strangers, thus generating 'automatic' control in an area without the use of vigilante groups, electronic supervision or simply locking strangers out, and so reducing certain street crimes.

The interaction and accessibility for strangers and inhabitants are profoundly influenced by the convex and axial organisation of an urban area and its interface with the buildings. The axial structure of the town allows strangers to enter an area, or, conversely, it keeps them out by making it difficult to get through (see Figure 7.9). The convex organisation of space and the interface with buildings – whether there are blank walls or barriers which distance the buildings from public urban space – may equally strongly affect the relation between inhabitants and their neighbours, and between inhabitants and strangers.

Deformed grids compared

These relations can be clarified by analysing the axial and convex maps of Apt (see Figures 7.9–7.13). The detailed plan of part of Apt shows how the islands of buildings form a system of open spaces which vary in width and length; the 'fatter' segments of space are knitted together by 'longer' segments, like beads on a string. The entire system of open space is lined with entrances to the buildings, creating a direct interface between the 'closed' and 'open' parts of residential areas.

The axial map of Apt (see Figure 7.11) shows how most of the central areas at the geometric heart of the town are easily accessible from the surroundings. Only one or two changes of direction – axial steps – are needed to reach the centre from outside the town. It is axially shallow, which is typical of many traditional towns, especially those whose existence depends on attracting visitors or passers-by to the commercial areas of the town. Furthermore, it can be seen that, although parts of the town are easily accessible, others are more difficult to reach – more axial steps are required.

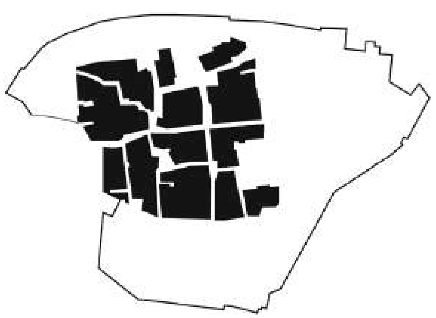


Figure 7.9 The area selected for analysis in Apt, a town in the south of France.

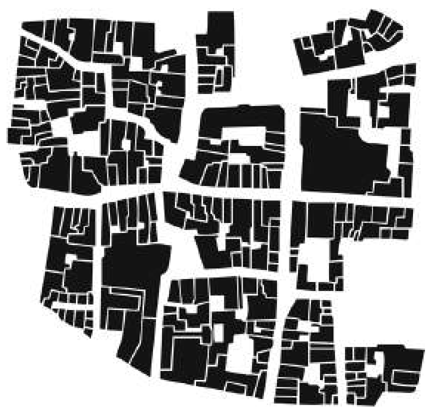


Figure 7.10 Block plan for area of Apt studied in detail.

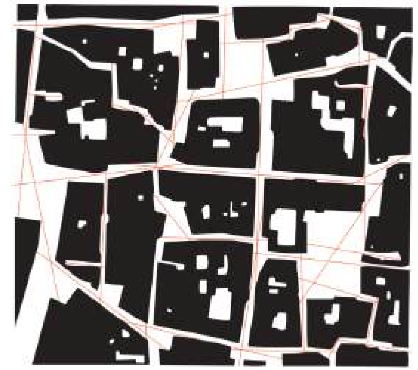


Figure 7.11 Axial map of an area of Apt.

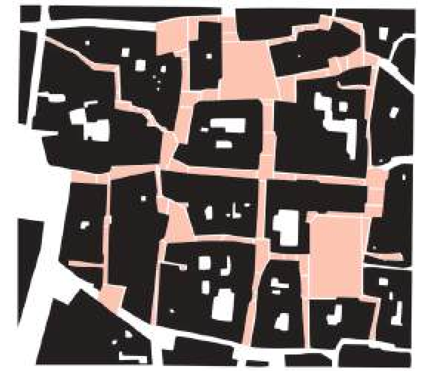


Figure 7.12 Convex map of the same area of Apt.

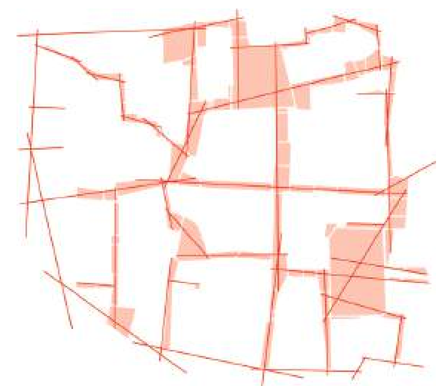


Figure 7.13 Axial and convex maps of Apt superimposed.

The convex map of this part of Apt shows the fewest and fattest segments covering this area (see Figures 7.10 and 7.12). The length and width of the convex segments and their variety are spatial characteristics of traditional towns. Building entrances open directly onto each convex segment of space. Only a few spaces are lined with blank walls.

When one superimposes the axial on the convex map, one can see that every axial line passes through several – often many – convex segments. This ensures local variation (the much-admired local articulation of vernacular space) but is also the means by which one is made aware of how the local parts are integrated into the town's global structure. This, coupled with the already noted way the building entrances relate to all convex spaces, contributes to the formation of that unique quality of urban space (see Figure 7.13).

Let us contrast this pattern of layout at Apt with a typical valiant attempt at recreating urban space, the Marquess Road estate in Islington. This estate, built in the mid-1970s as local authority housing, was intended to remedy the urban desert effect of many 1960s high-rise schemes. Our everyday experience of that estate and the systematic observation of patterns of occupancy confirmed that it differed far less than expected from the hostile environments it sought to replace. It was chronically under-occupied. In contrast to the spatial qualities of Apt, those of the Marquess Road estate emanate from Oscar Newman's theories of defensible space. Space syntax seriously questions the basis of Newman's theories.

At first glance, the plan of the Marquess Road estate in Islington seems to reproduce the informal arrangement of a vernacular village, with narrow streets, paths and village squares. Careful 'syntactic' analysis of the Marquess Road estate shows that, what first seemed a characteristically traditional morphological pattern, in fact contradicts the basic syntactic principles by which space constructs relationships between inhabitants and between inhabitants and visitors.

Both examples are deformed grids, but the greater deformation at the Marquess Road estate leads to a spatial pattern with fundamentally different potential social characteristics. The over-fragmentation of the Marquess Road estate has produced a labyrinthine rabbit warren rather than an urban village. (See Figures 7.14 and 7.15.)

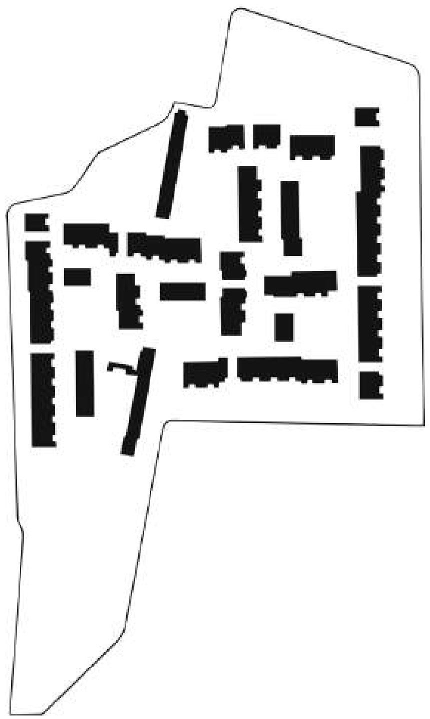


Figure 7.14 Plan of the Marquess Road estate.

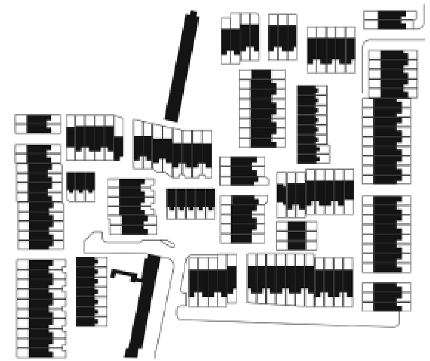


Figure 7.15 Block plan for area of the Marquess Road estate studied in detail.

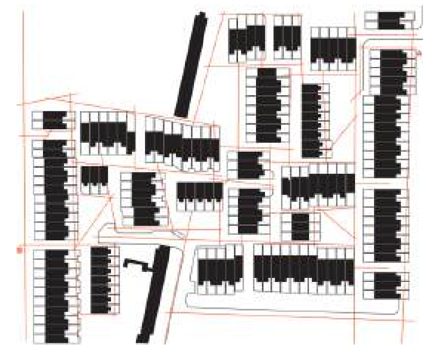


Figure 7.16 Axial map of an area of the Marquess Road estate.

The axial map of the estate shown in Figure 7.16 highlights one of the most common misconceptions about traditional urban forms found among contemporary designers: the axial organisation is very fragmented, with a high percentage of short axial lines placed orthogonally to each other. There are no axial lines to allow one to penetrate the scheme's geometric centre from the surroundings. The result is a system which is axially deep from the outside and other parts of the estate. To reach point B from point A – a similar distance to the same journey in Apt – requires at least eight axial steps at the Marquess Road estate, thus increasing the overall depth of the system. This makes it more difficult for strangers – or other people living on the estate – to pass through, thus increasing the segregation of the scheme from its surroundings. Rather than creating a cosy and intimate environment, we argue that the organisation of space has contributed to its isolation and desolation.

The convex map in Figure 7.17 reveals equally difficult problems: the front doors of the houses are concentrated around a few convex spaces, making it hard for inhabitants to recognise where their neighbours live. When one superimposes the convex on the axial map in Figure 7.18, it appears that the convex spaces, rather than being threaded along an axial line, are often as long as the axial lines themselves, so that the intelligibility of the whole is not evident at the local level. This peculiar spatial property may explain why so many people on the estate seem to be lost, and why there are many small squares and openings which are difficult to reach from any other part of the scheme. It may also account for the sudden appearance of people (to the great surprise of the inhabitants) due to the axial fragmentation of the layout.

Figure 7.17 Convex map of the same area of the Marquess Road estate.

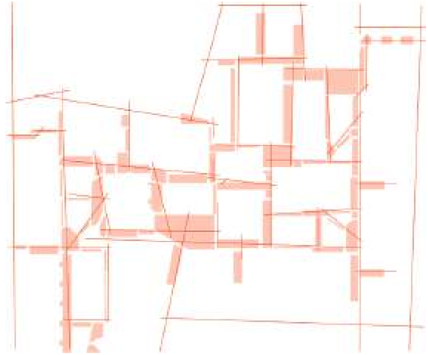


Figure 7.18 Axial and convex maps of the Marquess Road estate superimposed.

Figure 7.19 Numerical coding of the axial map of Apt.

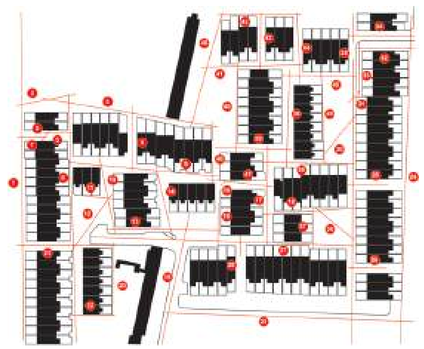


Figure 7.20 Numerical coding of the axial map of the Marquess Road estate.

Testing the analysis

To understand properly the relation between space and people, and in order to be able to test our hypotheses, we need more powerful and precise tools. The concepts we have introduced can lead to simple ways of obtaining certain key properties of space in numerical form. (See Figures 7.19–7.23.)

Pattern of depth

The concept of depth is one of the most important relational properties in space syntax. Depth exists where it is necessary to pass through a number of intervening spaces – axial or convex – in order to reach a space. Shallowness is where the route between spaces is more direct.

To clarify the point that depth can be seen and experienced from any given point inside or outside a system of spaces, the axial (or convex) map can be turned into a graph. The particular system can then be represented as a depth diagram, or justified graph, from any chosen point, for example from the town square or a peripheral street.[6] This difference can be computed numerically to give a precise index of relative depth or shallowness of any spatial system as seen from one particular point (see Appendix).

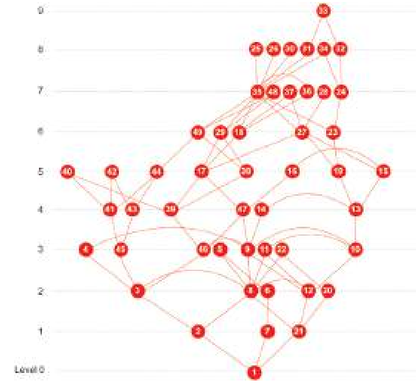


Figure 7.21 Depth diagram drawn from the entrance point to the Marquess Road estate.

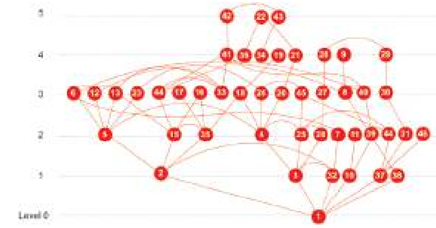


Figure 7.22 Depth diagram drawn from the market area of Apt.

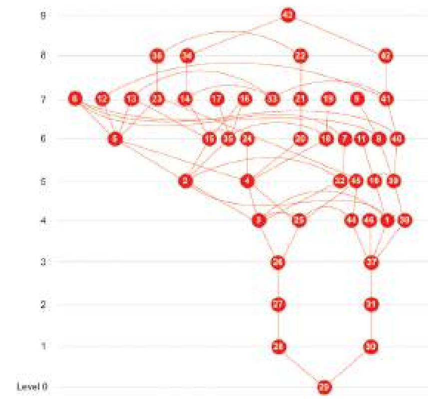


Figure 7.23 Depth diagram drawn from a residential street in Apt.

If one numerically codes all the lines on the axial map of Apt, a justified graph can be drawn using a computer (Figures 7.22 and 7.23). Any point on an axial line can be taken. The graphs are drawn by placing the point of origin at the baseline of the graph; all the spaces (in this case axial lines) one axial step deep from those points are placed one line above at depth level 1; all spaces two axial steps deep are placed two lines above, and so on until all the spaces in the town have been included. Then the depth structures of different towns and urban areas can be compared.

If different urban areas can be compared graphically and numerically, then the same system can be analysed from different spaces within it. This will indicate an often unexpectedly large difference, in terms of depth structure, in the same system seen from different points. For example, compare depth from two different points within Apt: drawn from a point in the market area, Apt seems very shallow (Figure 7.22), while from another point in a residential area the same system seems deep (Figure 7.23) – this is captured by a higher relative depth (RD=1.48) in the second case compared to the first (RD=056). The depth diagram from the entrance point to the Marquess Road estate shows it to be very deep (Figure 7.21).

The crucial point is that every space can be assigned a value that characterises its relation to all other spaces in the system, thus providing a global index for each space. The mean of all these depth values will then represent the depth or shallowness of the system as a whole, so it is possible to compare the depth patterns of different towns and urban areas.

Integration and segregation

Only by numerically comparing the values of different spaces can we reveal one of the fundamental properties of the spatial organisation of towns and urban areas: their structure of integration. Towns are arranged in such a way as to give priority to certain spaces: the main square or common High Street will tend to be shallower and thus more generally accessible than more secluded, deeper, quiet areas. Major commercial and public facilities will be within easy reach of other parts of the town. Although nowadays we can, to some extent, overcome both metric and syntactic distance using modern transport, our research suggests that the key difference in urban quality between spaces lies partly in how well the spaces are integrated with their surroundings.

By analysing the relative depth and ascribing a numerical value to each space, we can get a picture of the relative integration of each space. An axial line that is shallow from all points will be integrated, whereas one that is deep will be segregated.

Evaluating integration

This will allow us not only to understand an area's existing structure by identifying its most integrated and segregated parts, but it will also permit a design proposal to be slotted into the existing site so we can evaluate its effect on the total integration pattern. Quite simply, if a new shopping street is proposed, it will be possible to calculate whether the space will be well integrated with its surroundings. Clearly, the more a space is integrated, the more it may be able to exploit the existing pattern of movement of people caused by the arrangement of space.

Figure 7.24 Axial map superimposed on map of Apt.

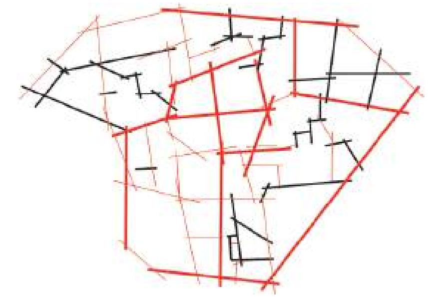


Figure 7.25 Axial map of Apt with integration analysis. Thick red lines are the most integrated. Black lines are the least integrated.

Once the relative depth or integration values have been worked out (either manually for small systems or by computer for larger ones), the results will show the distribution of the most integrated or segregated spaces in a town (see Figures 7.24–7.25).

The set of the most integrated axial lines, the 10 per cent most integrated lines in Apt (the integration core), can be drawn onto the map of axial lines (in thick red lines, see Figure 7.25). The percentage of spaces that are selected to define such a core varies according to size. The 10 per cent most integrated axial lines clearly pick out the underlying structure of the global integration, whereas in larger urban areas it is enough to look at the 5 per cent most integrated axial lines. The set of the most segregated spaces – in this case 50 per cent of the total – is selected from the computer printout and its distribution across the urban area will indicate the position of the segregated, or most inaccessible, areas in the town (see thick black lines). With this knowledge, we can evaluate any adjustments or additions to the town. The more a space is integrated, the more it may be able to exploit the existing pattern of population movements.

Let us once more compare Apt and the Marquess Road estate. Apt has an integrated core that forms a cross-like pattern which traverses the town, linking the centre to the outside (Figure 7.25). This confirms the global nature of the spatial arrangement because the integrated core links the commercial central streets and squares to the outside, making access easy. The segregated areas tend to form continuous clusters of spaces in the four quarters or interstices of the integration core. These spaces are in fact the residential quarters of Apt.

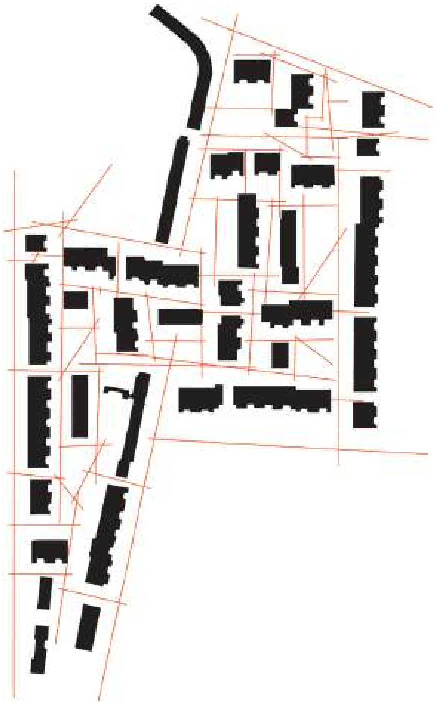


Figure 7.26 Axial map superimposed on map of Marquess Road estate.

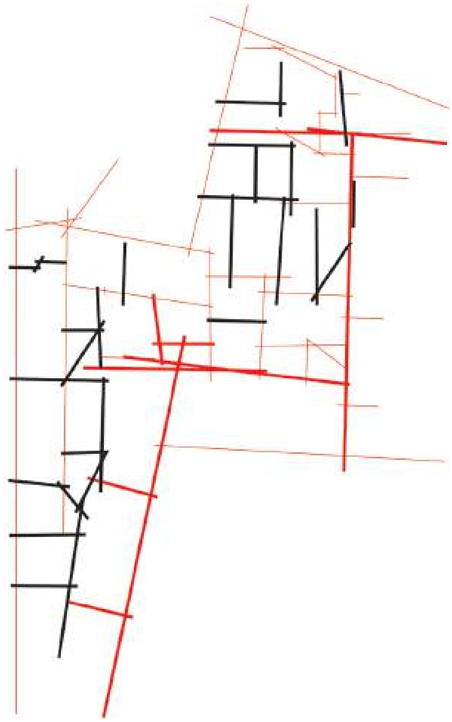


Figure 7.27 Axial map of Marquess Road estate visualising integration analysis. Thick red lines are the most integrated. Black lines are the least integrated.

The delicate balance between the position of the globally oriented spaces (the integration core) containing shops, offices, etc. and the local residential areas (the segregated zones) creates the subtle variation in spatial intensity of traditional towns: there is none of the crude distinctions into semi-private, semi-public and public space found in modern cities. Instead, the 'social' differentiation is achieved through the arrangement of space.

At Marquess Road estate the integrated core is not continuous and is concentrated on the edges of the estate rather than linking the central parts to the outside. This accounts for another finding from our observations on the estate, namely that most people were seen on the peripheral areas, and also explains the desolation of most of the inner areas which are generally very segregated. The estate's fragmented layout – which results from the enclave philosophy – produces large inaccessible spaces that are distant from the integrated spaces in the core, thus creating a fortress in the urban landscape (see Figures 7.26–7.27).

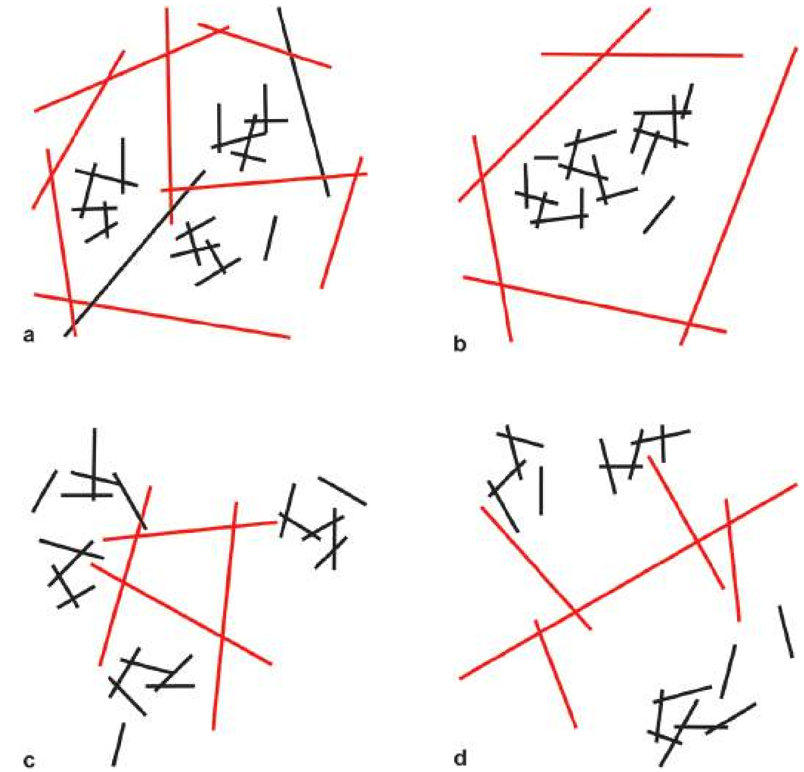


Figure 7.28 Integration core theoretical examples: (a), (b), (c), (d).

Any spatial pattern will produce a set of the most integrated or segregated spaces. A perfect grid will yield the same values for every space, while a deformed grid will produce a non-uniform distribution that distinguishes the different parts of a town. It is how these properties are distributed, and how they relate to the convex organisation and the interface, that must be understood in investigating the structure of a given area. (See Figures 7.28a–d.)

A theoretical typology of integration cores is presented in Figure 7.28. Figure 7.28a is an example of a globally integrated core: the most integrated spaces (in red) connect the centre to the outside, forming a wheel-like pattern. The most segregated spaces (in black) form clusters which are at the interstices of the wheel. In most types of traditional towns, it is the most integrated spaces that are lined by the major commercial facilities and pass through the major convex spaces or town squares in order to maximise the movement of people across the town. The segregated areas tend to be quiet residential areas, which are not hierarchically separated from the busy public spaces but are always near the 'integrated core'.

In contrast, Figure 7.28b is an example of an urban area where the most integrated spaces are distributed around the edges and do not penetrate the area's geometric heart. Conversely, the most segregated spaces form an inaccessible core at the centre.

Figure 7.28c is an example of an area where the integrated core is tightly wrapped around the centre, creating an inward-looking integrated heart which is inaccessible from outside. The segregated spaces form a band around the edge of the area.

Lastly, Figure 7.28d is an example of an area which has an integrated core that forms a tree-like pattern across the town – thus linking the centre to the outside – but leaving two large zones of inaccessible segregated spaces on either side of the integrated core.

Local control

Although integration can represent the global relation of a space, the local syntactic properties can be found by considering the relations between a space and its neighbours. The relationship between these local and global measures – as important in theory as they are in practice – has allowed us to construct a theoretical and suggestive model of urban space that defines the key spatial characteristics of intelligibility, continuity of occupation and predictability (see Appendix).

The Limehouse case study

With space syntax it is possible to take a new approach to urban design. By understanding the structure of the organisation of the area surrounding a particular site, valuable clues will be given to the 'spatial nature' of any proposed scheme. An architect may or may not want to integrate a new scheme with the surroundings and take advantage of the existing patterns of movement and occupancy. However, space syntax allows a proposal to be set in the existing urban fabric so that people can judge the effect of the scheme on the whole area and vice versa. This new approach, which can be tested and evaluated, we call 'global design'.

Space syntax in action

The controversial inquiry into the development of the Limehouse Basin in London's docklands provides a good case study (AJ 21.4.82) because there are several proposals that claim to contribute to the social integration of the area with the existing community.[7] The variety of formal solutions and social arguments lend themselves to precise analysis, as follows.

The area from West India Dock Road to the east, Salmon Lane to the north, Cable Street to the west and the Thames to the south was analysed and simulated. The axial structures of three design proposals, by R. Seifert & Partners, Shepheard, Epstein & Hunter, and the Limehouse Development Group,

were then analysed. Each 'network' was then set in the existing area to evaluate the syntactic properties and to simulate the patterns of movement produced by the schemes. The integration cores of the four systems (the existing system and the three proposals) were computed and graphically compared. The results can be summarised as follows.

The area as it stands is dominated too much by its peripheral spaces (the Thames, the Limehouse Basin and Commercial Road) and lacks both a local focus and well-structured patterns of movement through the area (see Figure 7.29). Too much space is used badly, and too much load is carried by too few peripheral spaces. The area's intelligibility is poor and has been made worse by much of the twentieth-century housing and by the decline of local industry and commerce.

Analysis of Limehouse now

Figure 7.29 shows that the area north of the Limehouse Basin is arranged in the traditional pattern of housing blocks and streets. There is a 'deformed' grid in the organisation of public open space with the buildings opening out directly onto it. In contrast, the area east of the basin contains different housing types – architectural experiments from the turn of the century to the 1970s – which invariably are isolated buildings surrounded by empty space (unlike the traditional street patterns to the north). The 1960s point blocks lack any urban structure at ground level, and the major commercial buildings are positioned away from the housing. Narrow Street to the south of the basin – the old Fore Street of the eighteenth century – still has some attractive buildings and suggests the scale and quality of the urban area before the 'destruction' caused by the housing interventions.

Figure 7.30 shows how the grid-like arrangement of space to the west and north of the basin becomes more fragmented as one goes east towards the housing between Narrow Street and Newell Street. There are few direct routes linking the

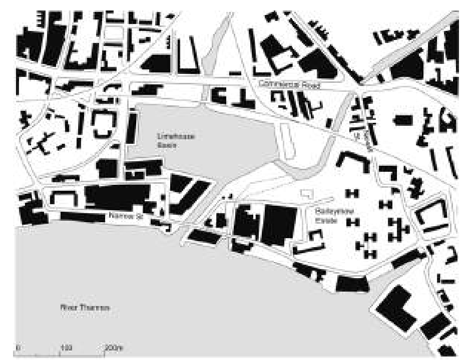


Figure 7.29 Diagrammatic layout: Limehouse Basin area as is.

Figure 7.30 Axial analysis of Limehouse Basin area as is.

area from north to south – from Salmon Lane commercial buildings to Narrow Street and the housing. Access to the waterfront is difficult, which results not only from existing property boundaries but also from the broken-up public roads and pathways. Today Limehouse is spatially dominated by Commercial Road. Most integrated spaces hang off Commercial Road and envelop – but do not surround – the basin, running from Commercial Road, Salmon Lane and Newell Street to Branch Road. These spaces tend to have the heaviest pedestrian (and vehicular) traffic. The south side of the basin and the area to the southeast, including much of the housing, is quite inaccessible compared with the whole area. This implies that a large continuous cluster of spaces at the heart of the area are cut off from their surroundings and would therefore tend to be underused or even unused. The computer simulation of movement through the area confirms that the edges of the site would tend to be heavily used while most of the central sections would be chronically under-occupied. From the above, it is clear that any scheme that tries to revitalise the area and reintegrate the basin with the existing community should, at least, take account of the distribution of the integration structure.

The best way to integrate the area would seem to be by providing a focus to the east of Limehouse Basin, linked to Narrow Street to the south and to the houses to the east – perhaps along the line of the existing viaduct. Without such amendments any development of the basin will remain isolated and contribute little or nothing to the surroundings.

Project analysis 1: R. Seifert & Partners

The main effect of Seifert's proposal (see Figures 7.31 and 7.32) will be to create a band of relatively segregated and inaccessible zones around the basin (see Figure 7.33). Figure 7.31, the block plan, reveals that the misleading, cosy, 'brick-and-trees' imagery of the plan is in danger of reproducing the unstructured arrangement of buildings surrounded by empty space which characterises the housing to the east.

The deformed grid of the whole area is radically changed by Seifert's intervention, as can be seen in Figure 7.33. The number of axial lines increases. The large number of short axial lines helps to create a rabbit warren rather than an urban area. Such an arrangement has been shown (in the Marquess Road estate analysis seen above) to increase disorientation and it functions as a barrier to informal journeys through the area. Hardly any spaces directly link the housing to the east with the waterfront; instead, the increased fragmentation of the spatial layout will make it more difficult – over and above property boundaries and barriers – to reach the basin and the facilities in the proposed scheme from the surroundings. The analysis confirms that Seifert's proposal does little or nothing to improve the overall integration of the basin in the area. In fact, although the integrated spaces are about the same as in the existing area, the scheme succeeds in creating a cluster of highly inaccessible, segregated spaces around the basin. The segregated areas have in fact shifted from the southeast of the basin to around the basin. The most inaccessible spaces of all are those along the

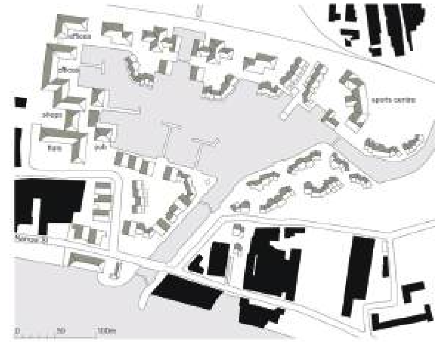


Figure 7.31 Proposed Limehouse Basin plan by R. Seifert and Partners.

Figure 7.32 Diagrammatic layout of Seifert's proposal in context.

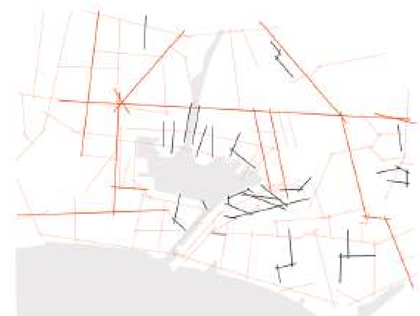


Figure 7.33 Axial analysis of Seifert's proposal.

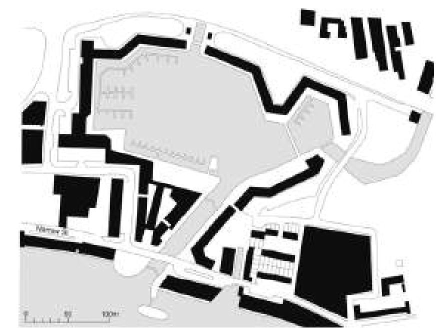


Figure 7.34 Proposed Limehouse Basin plan by Shepheard, Epstein & Hunter.

basin's eastern edge, which invalidates Seifert's claim that the water will become part of the public domain. The analysis reveals that Seifert's proposal turns its back on the community by introducing an alien layout.

Project analysis 2: Shepheard, Epstein & Hunter

In spite of appearing more like a globally conceived street system than Seifert's pseudo-village, Shepheard, Epstein & Hunter's scheme (see Figures 7.34 and 7.35) will have similar effects. The analysis shows that the scheme has tried to recreate a street-like arrangement with linear routes and more widely distributed facilities (see Figure 7.36).

Figure 7.35 Diagrammatic layout of Shepheard, Epstein & Hunter's proposal in context.

Figure 7.36 Axial analysis of Shepheard, Epstein & Hunter's proposal.

The block plan in Figure 7.35 reveals how the scheme does not reproduce the traditional pattern of the surrounding urban fabric but, as with Seifert's scheme, reads as an effective intrusion of the area. Access to the scheme is limited, making it difficult to reach the waterfront. The proposed scheme shows a preoccupation with the 'local' organisation of space which does not take account of the global spatial properties of the area.

The axial map in Figure 7.36 confirms how the scheme interrupts the grid-like arrangement of the surroundings. This island of snaking streets does not relate to the housing to the east, and nearly all connections to the north are severed. The southern part of the basin is better connected near the Limehouse Gallery, while much of the northern and western areas of the basin are difficult to reach from the surroundings. As a result, a large part of the scheme to the north of the basin is quite inaccessible. The southern part of the basin and the southeast waterfront become integrated, but there is a continuous band of segregated spaces which will produce the same ghetto effect as Seifert's scheme. The activities at the northern and western edges of the basin will tend to be more introverted and exclusive and will not serve the people living to the east. The analysis shows that this scheme does not succeed in integrating the basin with its surroundings.

Project analysis 3: Limehouse Development Group

The Limehouse Development Group's scheme centres on a new local market to the east of the basin, with direct links to Narrow Street by a new road. The analysis of the Limehouse Development Group's proposal in Figure 7.37 demonstrates that Limehouse can be improved and that the two other schemes do not exhaust the range of architectural possibilities for the area.

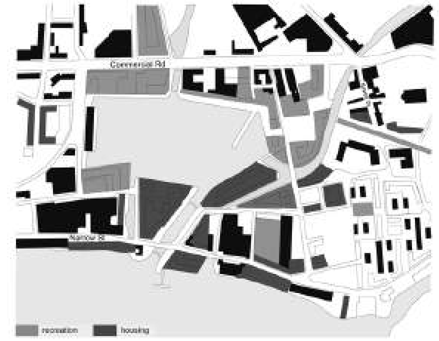


Figure 7.37 Proposed Limehouse Basin plan by Limehouse Development Group.

Figure 7.38 Diagrammatic layout of Limehouse Development Group's proposal in context.

Figure 7.39 Axial analysis of Limehouse Development Group's proposal.

The diagrammatic layout of Figure 7.38 links Limehouse Basin to Narrow Street and the Thames to the south, and reconstitutes the urban fabric in and around the housing areas. The sketch scheme also incorporates a market to the northeast of the basin which acts as a focus for the community to the east and the north. The basin waterfront is kept as accessible as possible so that it can be used by the existing community.

The axial map of Figure 7.39 shows that the number of axial lines has been reduced, and that many of the existing streets have been extended to link housing and the southern part of the site to the basin and the market. The axial structure of the proposed scheme reproduces the general spatial properties of the surroundings so that the new proposals do not read as an interruption but as a continuation of the existing urban fabric. The integrated spaces extend from Commercial Road to

Narrow Street, passing through the proposed market and the basin waterfront. The housing becomes integrated with the area and does not form large inaccessible and unpleasant spaces. Areas that are intended to maximise the flow of people – the commercial market and shopping streets – will be busier because of the spatial arrangement. Conversely, the office and upmarket housing to the north and west of the basin will remain relatively quiet and separated.

In general, the Limehouse Development Group's proposal creates a positive structure for the area which should play an important role in using everyday patterns of movement to bring life to a dead urban area. At a local level there are still a few problems because some important spaces remain over-segregated and under-used. Further development of this scheme ought to eliminate these problems since the global organisation of space has been well solved (see Figures 7.37–7.39).

Conclusion

In conclusion, the numerical and graphic analyses demonstrate an important principle: adding new compact development to the middle of an area will not improve its integration with the surroundings. With all its problems, Limehouse is better integrated with the undeveloped basin than with Siefert & Partners' or Shepheard, Epstein & Hunter's proposed scheme.

Epilogue

So how much does all this matter? Surely everybody now believes what sociologists have been telling architects for years: that space does not matter much because social relations are made by people, not by architecture, and they will do what they want regardless of design.

Our answer to all that is in three parts. First, we believe that the unforced presence of people (that is, they are not attracted by crude social magnets) that can be affected by spatial layout is sociologically important, but that sociology does not yet present us with concepts that allow this to be formulated easily. We believe that architects' desire and concern to see that the urban spaces they create are fully occupied and simply well used is a worthwhile social objective. The very sensation of people around you is, we believe, just as important to successful city life as are the personal relationships which occur in most cases independently of, and sometimes because of, particular juxtapositions and forms of public spaces and buildings.

Second, we believe that feeling safe in a city depends largely on areas being in continuous occupation and use – the feeling that, in even mildly populated urban or suburban areas, one is never walking around alone.

Third, we believe that there is a subtle relationship between the organisation of urban space and the types of society that occupy them, and in some way this relationship is explained by the very opposite of the territoriality theory. That theory assumes that individuals need to express themselves spatially through privacy; towns are explained as hierarchies of groupings, expressed in discrete and relatively inaccessible territories. On the contrary, the theory and practice of space syntax demonstrate that towns seem to be explicable in their form mainly as spatial structures that work against society: what society separates, space puts back together again.[8]

Space syntax is therefore both a method and a message, and it would seem to open up new perspectives to urban design. It gives a rational way of approaching urban design 'top down', so that anyone can participate in the decision-making process from the 'bottom up'. Space syntax allows the structure of the area to suggest new possibilities. Above all, it is a way of looking at the oldest problem of all in urban design: how to add the new to the old.

Space syntax: Appendix

Intelligibility

A system's intelligibility is defined as the correlation between global *integration* and local *control*.[9] Intuitively this means that the large-scale structure of a system is intelligible to the people moving about in it to the extent that the information they receive about the space they are in – the local connectivity and control – also allows them to comprehend the structure of the whole. This seems to capture the way people can learn about large patterns from their experience of small parts, or fail to do so when the correlation is weak. Typical urban areas or towns will tend to have an intelligibility correlation of about 0.45, while unintelligible systems will have values of 0.2 or even less, where a value of 1 is a perfect correlation and 0 is random.

Predictability

Global predictability is the correlation between the integration value of a space and its simulated or observed density of use and movement. By taking the average correlation for all spaces, we have a figure which characterises the overall predictability of an area or town. This confirms the everyday intuition that some spaces are quiet and others busy because of their position in the overall pattern, and this should be expected to match up with the position of different facilities. The average global predictability of the sample of urban systems is 0.75; in other words, a high correlation between integration and movement which has been confirmed by the observations of 'working' urban environments. An urban area in London has an average predictability of 0.8, while less predictable systems will be below 0.6.

Local predictability will be the correlation between the control values of spaces and the density of use and movement. An effective system of local predictability will have correlations of about 0.7, which means that large patterns of occupancy in an area can be inferred from the visible part of the system the individual is in.

Continuity of occupation

Continuity in an urban structure will then be indexed by the statistical bias of the distribution of people in space. A highly biased or structured system will be one in which too few spaces attract too many people, while too many spaces remain empty or have too few people in them to contribute significantly to urban life: it is this condition which characterises so many contemporary urban environments.

In general, urban areas with globally orientated integrating cores – which link the inner areas to the surroundings – will tend to perform well on all the above counts. But the core must not be too strong or it will attract too much movement at the expense of other parts and therefore interrupt the principle of continuous occupation. Too few over-strong spaces can drain the life out of the rest of the system and concentrate movement on a limited number of spaces.

The principle of the strong but not over-strong core seems well confirmed by observation. A more extensive programme of simulation studies has suggested a much more precise theoretical model of urban space and how it works. Theoretically, it turns out that local predictability in an area is strongly influenced by the overall integration of the system. The more segregated the area is, the weaker the local predictability – it will be more difficult to read the encounter potential of an area from its constituent parts and the overall pattern of encounter will seem more puzzling.

The area's global predictability, however, is influenced more by the strength of the correlation between local and global properties – what we have called intelligibility. Computer simulation has shown that not only does this crucial relation between the local and global structure of an area influence the psychological intelligibility of someone moving about, but it also plays a key role in structuring movement – even in a computer which has no conception of intelligibility.

Relative depth

The formula for calculating the relative depth (RD) is $RD = 2\frac{(MD-1)}{(k-2)}$, where *MD* is the mean depth of a system (obtained by dividing the total depth by k, the number of spaces in the system). This value of relative depth can be adjusted to eliminate the effect of size, giving a figure which varies about 1 (the average for a cross-section of urban systems), with low values indicating shallow or integrated systems, and higher values for deep or segregated systems, allowing comparisons between towns and urban areas.

Control value

The control value is calculated by ascribing a value of 1 to each space, or axial line, then summing the reciprocal of the valences of neighbouring spaces, which will give values over 1 for spaces more connected than their neighbours and values less than 1 for less connected spaces.

Notes

1. Sir Norman Foster, 'Opening address', in *First International Space Syntax Symposium, April 1997*, edited by M. D. Major, L. Amorim and F. Dufaux (London: University College London), pp. i–vi.
2. B. Gil and C. Coelho, 'Laying the fundamentals: Early methods and intentions from the outsets of space syntax', *Proceedings of the 11th International Space Syntax Symposium* (Lisbon: Instituto Superior Técnico, 2017), 13.1–13.13.
3. The SERC, the UK Science and Engineering Research Council, was the principal source of funding for the Unit for Architectural Studies until the early 1980s. The Unit has undergone several reincarnations since then and is now known as the Space Syntax Laboratory (eds).
4. The original photographs by Dougie Firth and accompanying illustrations by Gordon Cullen have been omitted due to copyright restrictions. The original article can be obtained from *The Architects' Journal (Archive: 1929–2005)*, 178(48): 47–54, 59–63. Retrieved from https://www.proquest.com/trade-journals/space-syntax/docview/1427086296/se-2.
5. The text is referring to the illustrations by Gordon Cullen in the original article. Please see the note above (eds).
6. The term 'justified graph' has become the definitive one in space syntax literature (eds).
7. The authors are citing an *Architects' Journal* Report: 'Hunting Gate amends Limehouse scheme', *The Architects' Journal* 175, no. 16 (1982): 42 (eds).
8. This key idea is also expressed in the 1976 'Space syntax' paper: 'Space is not a reflection of society, but a set of strategies in relation to social form, as often as not offering an alternative basis for encounters other than those dictated by the social structure' (see Chapter 6 of this book). But it found a powerful and more crystallised expression in a later article (J. Hanson and B. Hillier, 'The architecture of community: Some new proposals on the social consequences of architectural and planning decisions', *Architecture et Comportement/Architecture and Behaviour* 3 (1987): 351–273, 265): 'Space may not be structured to correspond to social groups, and by implication to separate them, but on the contrary to create encounters among those whom the structures of social categories divide from each other. In other words, space can in principle also be structured, and play an important role in social relations by working against the tendency of social categorisation to divide society into discrete groups. Space can also reassemble what society divides.'
9. Editors' note: the definitions throughout this Appendix should be read within the context of the article. The terms were developed through the research undertaken at the Unit for Architectural Studies – which became the Space Syntax Lab in later years – in the early 1980s, which used software developed by Paul Coates and later by Nick Sheep Dalton. The definitions cannot be transferred to segment analysis, which has become more common in subsequent work.

8 What do we mean by building function? (1984)

Introduction to 'What do we mean by building function?'

Philip Steadman

This conference paper was given just before the publication of *The social logic of space* and presents some of the key ideas of that book in highly abbreviated form. The authors call for a 'theory of description' of architectural form and function, and regret the absence of efforts to address this issue in contemporary architectural research. Some of the basic contents of the space syntax toolbox are introduced. The 'justification' of graphs of access or 'permeability' is explained with the example of a house converted into a doctors' surgery. The concept of an 'architectural genotype' is introduced with a simple seventeenth-century house plan from Wood-Jones's Banbury sample.

The final section of the paper discusses in more informal terms the idea of 'global function', meaning the 'schoolness' of a school or the 'churchness' of a church. The conceptual distinction is introduced between the 'occupants' of a building and 'strangers' who visit it, with interfaces created to separate these groups. A doctors' surgery is taken as an example. This has a very different organisation in terms of access from the house from which it was converted. A similar organisational structure to the surgery is replicated several times in a health centre, but not in two hospital outpatients' departments, in which a 'morphological discontinuity' is diagnosed. Larger buildings present perplexing problems for the theory of description.

Several points strike me, as a relative outsider then and now, but one who has followed the evolution of the space syntax programme closely since its inception. I am surprised, as an alumnus of the Centre for Land Use and Built Form Studies, set up in 1967 (working on graph theory and plan layout since 1973) to learn that research in architecture had up to the 1980s 'studiously avoid[ed] the issue of form'. This perhaps, however, was a shot directed at the sociological and 'building utilisation' focus of the conference to which the paper was given. More important is the claim, both in the text and implicitly in the title, that a general definition of building function and its relation to form could be based wholly on a topological description of the patterns of access between spaces in buildings. The truth is that graph representations actually discard much of the geometry of room layout and built form.

The paper does not cite directly what in my assessment is one of the most interesting and suggestive early products of the Bartlett group, *The analysis of complex buildings* by Hillier and Hanson together with Paul Stansall, of 1978.[1] But it picks up the thread. This internal report, now hard to find, presents the plans of 24 large buildings of different types and 'global functions', including one hospital, with their justified access graphs. Values for a series of measures on those graphs are tabulated. The purpose is to test the graph representation, and to move towards the identification of genotypes that would characterise 'schoolness', 'churchiness' and so on. In the event this research programme, which in my view held great promise and interest, was not pursued; and attention turned from access relations to pedestrian movement and visibility in large open architectural interiors.

What was the reason for this programme being abandoned? There are hints in the report itself. In large buildings, basic spatial layouts can be repeated on successive floors, constrained by the patterns of vertical circulation. The geometry of the horizontal circulation begins to be governed by considerations of visibility and legibility. Modular plan elements consisting of groups of rooms are repeated. As the authors say: 'It may be that many buildings are unique global concatenations of well-ordered sub-complexes.' It might not therefore be feasible to identify genotypes for entire large buildings; but this might be possible for the sub-complexes. One cannot help feeling that the use of the word 'complex' in the report's title signals an anxiety that these buildings are just *too complex*.

Adrian Forty, in his historical account of how the word 'function' has been used in architectural writing, stresses how up until the twentieth century it referred almost exclusively to tectonic or structural constraints on built form, and then was extended to refer to the environmental functions of buildings and how they control the weather and admit daylight.[2] It seems to me that the difficulties that space syntax came up against at this period, with the analysis of complex buildings, have to do with an overly narrow definition of function and its relation to form as advanced in this paper.

The access structures of buildings, especially as they become larger, are constrained by Forty's other classes of building function: by the limits of structural spans, by the depth in plan to which daylight will penetrate, by circulation becoming organised into long straight corridors. It is not just social functions that are expressed in graph structure. Maybe a newly revived programme of *The analysis of complex buildings* could bring these different classes of function together and find a wider scientific definition of the relationship of architectural form to function. I have been thinking about this.

What do we mean by building function?

Bill Hillier, Julienne Hanson and John Peponis

Originally published: Hillier, B., Hanson, J. and Peponis, J. 1984. 'What do we mean by building function?' In: Powell, J., Cooper, I. and Lera, S. (eds.), *Designing for building utilisation*. London: Spon, 61–98.

Scientific approaches to architecture usually avoid the issue of building form, preferring to focus on function. But how can there be a theory of function without a systematic analysis of the key architectural variable of form? A theory of description is required. In this paper it is argued that such a theory can be built through the analysis of spatial form in buildings. Then once spatial form is describable in terms of a descriptive theory, a more powerfully scientific – and architectural – understanding of function is possible. The argument draws on several pieces of research carried out by the authors and their students, but focuses eventually on various types of medical building in order to illustrate certain general principles.

The need for a theory of description

There are, it seems, now two distinct traditions in architectural discourse: a *critical* tradition, which is concerned first and foremost with the changing *form* of buildings, sometimes confining itself to the superficies of style, but at its best attending carefully to the systematics of spatial and morphological form (the work of Colin Rowe and Paul Frankl comes immediately to mind;[3] and a *research* tradition which (as witness the preoccupations of this conference) studiously avoids the issue of form, and addresses itself almost exclusively to matters of *function*, in the belief, it would appear, that function is scientifically tractable whereas form is not.

But surely the schism is bizarre. How can there be a useful theory of building function unless it either incorporates or relates to a theory of the architectural malleability of form? Similarly, how can there be a theory of architectural form independent of a theory of the functional logistics of form? For both the scientist and the architect it is the essence of his discipline that the two issues are aspects of a single question: what is it about architectural form that *works*?

For the architectural researcher, the question ought to be crystallised as his or her most pressing concern, since how can any investigation be truly systematic unless the architectural variable can be controlled? That architectural variable is building form. Architecture decides form, and hopes for function. Controlling the architectural variable for research purposes means having a *theory of description* of building form. Without it, we can do nothing. With it, we can at least begin to find out what there might be to know about buildings.

The aim of this paper is to set out a theory for the description of building form, and through this to arrive at a reformulation of the problem of function. More precisely, it tries to answer two questions:

- What is it about what people do (function) that leaves its mark on building form? and;
- What is it about building form that leaves its mark on what people do?

Both of these questions are, we believe, capable of exact answers given a theory of description.

Elements of a spatial description

First, we must decide what it is about buildings that needs to be described. At first, this appears an impossibly complex question, not the least because buildings are so many different things: they are physical constructions, they are arrangements of space and they are objects in a style culture. But provided we remember that only art deals with the whole of reality and science can only hope to understand its underlying dimensions, then we need not be pessimistic. Being both physical objects of some use value and also stylistic objects is a duality that buildings share with most other artefacts.

It is the organisation of space that makes the building unique. It is the distinguishing mark of space that sets the work of architecture apart from other artefacts, and it does so in a very important way.

A bridge, a vase or a surgical instrument is functionally useful insofar as it is a physical construction, and socially meaningful insofar as it participates in a style culture. Physical form and social meaning are, to some extent, separable questions. In architecture, this is not so. Because buildings organise space for social purposes, social considerations are present in the very physical form of the building. Social meaning is not a gloss added to buildings: it is an intrinsic aspect of their physical form. Indeed, it is their most important aspect. Buildings are *for* the organisation of space. This is the most general statement of function. In buildings, technology permits, and style confirms, that space has been created and organised for social purposes. This is the unavoidable basis of our discipline.

The task of a theory of description thus becomes one of a theory of spatial description. But not only that. If spatial description is to reveal functional dynamics, then it must somehow incorporate or lead to a theory of the social nature of space. We must learn to describe buildings spatially as social products, or we will not be able to arrive at a usable definition of building function.

A descriptive theory for space has, in our view, to be built at three levels:

1. the identification and representation of spatial *elements*;
2. the categorisation and analysis of spatial *relations*; and
3. the modelling of common or '*genotypical*' *themes* and patterns.

The first, identification and representation of elements, must, we believe, be solved in three ways, not one. The first identifies *boundaries*. A building is a more or less controlled domain, and this means a continuous boundary perforated by one or more entrances, possibly with other boundary-entrance pairs within, as can be seen in Figures 8.1a and 8.1b.

Figure 8.1a is a simple plan and Figure 8.1b represents its boundary structure as a graph in which circles represent bounded spaces and lines relations of direct permeability.

But bounded space by no means exhausts what a building is spatially. Some very complex buildings have very few bounded spaces – churches, some schools, offices and so on – yet achieve a highly complex and differentiated pattern of space. These can be analysed – at the level of the plan at least – by introducing two further kinds of spatial description. The first we call a 'convex', or two-dimensional

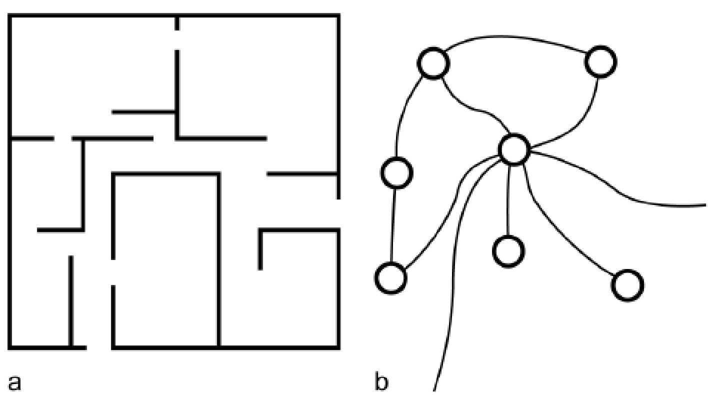


Figure 8.1 (a) Plan and (b) boundary graph.

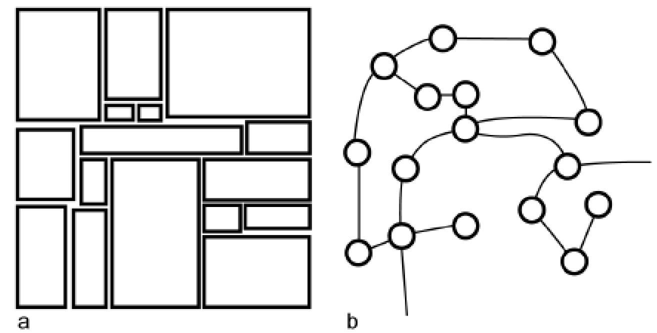


Figure 8.2 (a) Convex break-up and (b) convex graph.

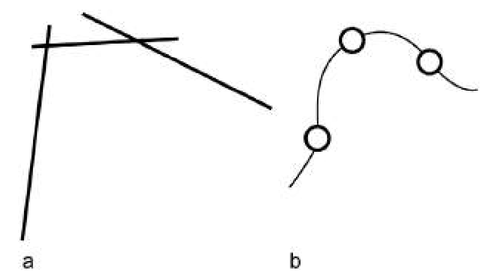


Figure 8.3 (a) Axial map of circulation and (b) axial graph of circulation.

description. This identifies the fewest and fattest convex spaces that cover the system, by applying a rule which says that fat spaces always prevail over thin spaces. Figures 8.2a and 8.2b show our example analysed convexly, together with its graph.

The second we call an 'axial' or one-dimensional description. This identifies the longest and fewest straight lines that cover all the convex spaces in the plan. Figures 8.3a and 8.3b show the circulation spaces of our example analysed axially, together with its graph.

Even before we begin a relational analysis, these simple representational devices can allow us to quantify and represent visually many of the spatial characteristics of plans, the more so as plans become complex and everyday language fails to provide a rigorous account. To begin with, we can simply count the numbers of bounded, convex and axial spaces and express them in relation to the total area of the plan – so many spaces, of one or other type, per so many square metres. These will tell us the *rate* at which a particular type of building or a particular architect adds each type of space to a growing structure.

More informatively, we can look at the ratios of each type of space to the others. Since the convex system must always have the maximum number of spaces, then it will serve as the best reference point. If we divide the number of bounded spaces by the number of convex spaces, then we arrive at a figure between 0 and 1, where figures close to 1 indicate few convex spaces over and above the bounded, and close to 0 indicate many more. Low figures thus indicate a higher degree of *convex articulation* of the plan.

If we then divide the number of axial lines by the number of convex spaces, then we will have a similar measure of *axial linking*, since if few lines link many convex spaces, we will have a low value, and if the number of lines approaches the number of convex spaces, we will get a value close to 1, and little axial linking. We can quickly establish that, for a good sample of his villas and apartments at least, Corb [Le Corbusier] convexly articulates more, and axially integrates less, than other major modern movement figures.

It should be noted, by the way, that the measure of axial linking only works if the fewest axial lines to cover the convex spaces have been drawn. In cases where extra links between convex spaces can be made by additional axial lines, then the best measure of the axial organisation is one which compares the axial system to a perfect orthogonal grid by using the following formula where R is the number of rings in the axial map and A the number of axial lines. High values indicate an axial organisation approximating a grid, very low values indicate a high degree of axial break-up. (Readers interested in the full account of these measures should consult Hillier and Hanson 1983, forthcoming.)[4]

$$\frac{(2\sqrt{R})+2}{A}$$

Relational analysis

These representations and figures give useful data on the general characteristics of plans. Much greater precision in analysis can, however, be achieved through the numerical analysis of relations among spaces. The *integration value* of a space (code-named RRA in the house jargon of the space syntax research programme) will express how many spaces distant a particular space is from every other space in the plan, it being the case (though far from always obvious) that these values are different from one space to another in the same system. Figures 8.4a and b show this graphically and numerically by taking the two convex spaces in our example, and 'justifying' the graphs from those points, that is, aligning all the other spaces above the selected space in layers according to how many spaces each is from it.

The numerical value is a number varying about 1, with low values indicating more integration (in effect, less distance to all the others) and high values more segregation, or more distance. It is calculated first by the following formula where *MD* is the 'mean depth', or the mean number of spaces away, of all other spaces, from the selected space, and *k* the total number of spaces in the system, and then applying a correcting factor to eliminate the empirical effects of size (see Hillier and Hanson for details)[5] and permit cross-comparisons of systems of different sizes.

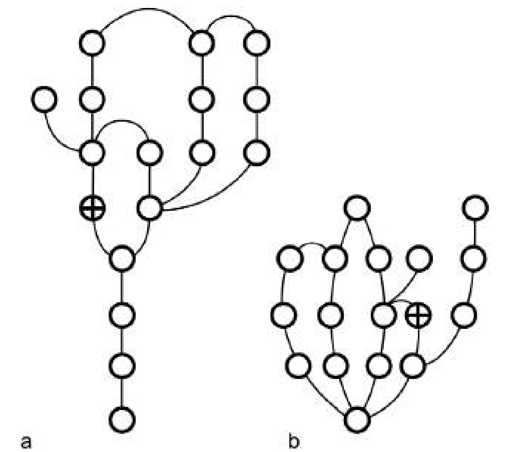


Figure 8.4 (a) Figure 8.2 justified from its most segregating space, RRA = 2.152; (b) Figure 8.2 justified from its most integrating space, RRA = 0.717.

$$\frac{2(MD-1)}{k-2}$$

The mean of all the integration values from spaces in a system will then precisely express the overall degree of integration or segregation in that system. The 10 per cent (or 2 per cent or 25 per cent) of most integrating spaces will then form the '*integration core*' of that system, and likewise for segregation. In real systems, integration values correlate highly with patterns of global movement, to an average of about 0.75.

Integration is a 'global' measure since it takes into account every other space in the system. 'Control value', on the other hand, is a 'local' value, since it takes into account only the neighbours of a space and the neighbours of those neighbours. It expresses, again with a value varying about 1 (but with high values indicating strong control), how much better or worse connected a space is than its neighbours. It is calculated simply by summing the reciprocals of the valencies (numbers of connections) of the neighbours of a particular space – see the formula below:

$$\sum\frac{1}{C_N}$$

In effect, strong control spaces gain more than they give away, and vice versa for weak control spaces. In general, control values correlate very highly with local patterns of movement, to an average of about 0.87.

By applying integration and control measures to the subset of spaces that lie on at least one ring in the system – that is, those which offer at least one alternative route to other spaces, in contrast to a ringless tree form where there is only one route from every space to every other space – one can then work out how a space relates to the route choices available in the system, that is, how far it integrates and controls those choices.

Genotypical themes and patterns

Differences in integration and control values from one space to another and from one part of the system to another, applied to the bounded, convex and axial patterns, are among the fundamental means by which social relations put their imprint on buildings. 'Differences' is the key word. Take the ordinary house, for example, so often repressively interpreted in recent years as a set of relationally identical function-spaces linked to a central corridor. In every vernacular tradition of housing we have examined (see Hillier and Hanson 1982), the house is a set of spaces which relate *differentially* to the overall spatial pattern of the house. Each function is, in effect, recognisable from its unique set of integration and control characteristics. Function is what people do – but it is also a characteristic spatial pattern, and a characteristic part of an overall spatial pattern.

Take an extremely simple example (from Wood-Jones).[6] Figure 8.5 shows a simple plan, and how different functions have different integration characteristics, both visually using the justified graph technique, and numerically using integration values. In fact, this pattern in which, of all the daytime spaces, the parlour (P) integrates least, the main living area (L) most, and the kitchen (K) lies in between is a very common pattern, and underlies many different housing geometries. We have come to call this type of relational differentiation of the functions and categories located in different spaces within a plan an 'inequality genotype', because these important spatial differences can be expressed as numerical inequalities.

Now if such 'inequality genotypes' can be shown to exist across a sample of buildings within a particular cultural tradition – and this can obviously be simply by aggregating, taking means, testing the strength of differences across a sample

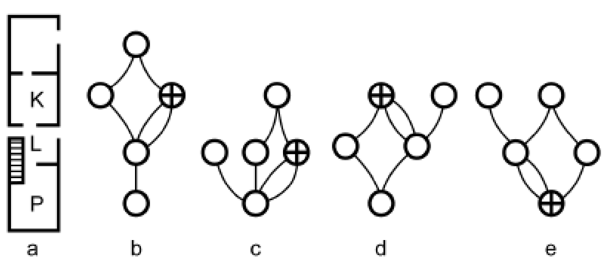


Figure 8.5 Plan of a seventeenth-century Banbury house and justified graphs from P, RRA = 0.667; L, RRA = 0.167; K, RRA = 0.333; exterior, RRA = 0.333 (left to right).

as compared to individual cases and so on – then it is reasonable to say not only that we have identified a cultural genotype by objective and numerical means, but also that we have done so by means of an analysis which is at once functional and spatial. We have simply looked at the relational spatialisation of those different functions. Once we had a theory of description, the problem in effect disappeared, and we were able to detect what we all know to exist anyway: the imprint of social relations and – dare we say it – even of the human mind on these extrasomatic organs we call buildings.

Global functions and interfaces

But what about the overall or 'global' pattern of space? For this we need a new concept which we will call 'global function'. 'Global function' means something like the overall figure that characterises different 'types' of building: the 'churchness' of a church, 'schoolness' of a school and so on – in effect, the link between spatial form and building function that everyday language affirms by referring to both with the same word.

All such typical figures have, it seems, certain things in common. All building 'types' define two fundamentally different categories of people: a set of '*inhabitants*' whose social identity as individuals is durably recorded in the building form by control of space or a set of spaces; and a set of '*visitors*' whose rights of presence in the building exist and distinguish them from the world of strangers, but not in a durable way as individuals and not through control of spaces. Family members in a dwelling, teachers in a school, medical professionals in a hospital are all inhabitants in this sense, while guests in a house, pupils in a school and patients in a hospital are all visitors. Note that the length and constancy of occupation are *not* the criterion of membership. In many institutional buildings – asylums, prisons and so on – visitors are much more permanently present than the inhabitants who escape at every opportunity.

In general, buildings can be defined in these terms as devices for making two kinds of *interface*: one between inhabitants and visitors, and the other between different categories of inhabitant. These interfaces are realised through some configuration of bounded, convex and axial space, and some set of relations of integration and control. These interfaces are what we mean by the global function of the building. It is this, we believe, that we name – at once socially and spatially – when we say 'school' or 'church' or 'hospital'.

It would seem then that if we could find a way to analyse and classify 'global functions' in spatial terms, then we might be on a way to a theory of building form which linked the traditional notions of 'type' and 'function' in a new and systematic way. Unfortunately, although such an analysis can be very rewarding, we do not believe it can be completed in quite those terms. There are two main reasons for this. First, there are paradoxes and difficulties in the way of realising all but the simplest interfaces in perfect spatial form. For example, the preservation of statuses often requires segregation, while the parallel requirement to control others requires close proximity, and these contrary pulls can produce as great discontinuities within a type of building as between that type and others. Second, the effect of size is such as to, in itself, radically alter possibilities and even the desirability of realising social structures in spatial form. Buildings literally become quite different objects socially as they become large, and set us problems which are more the problems of large buildings in general than the problems of a particular building type.

What follows, while sketchy, draws on several studies by ourselves, colleagues and students involving spatial analysis to identify genotypical patterns, field observations of space occupancy and movement, and computer simulation experiments on real buildings to see how space affects function theoretically. For the sake of continuity and clarity we will concentrate on 'medical interface' buildings, although our real purpose is to show general principles.

Medical interfaces and others

The simplest 'medical interface' is probably the common-or-garden doctors' surgery, often adapted from an ordinary house. It turns out that our initial example was such a building. Figure 8.6 shows the plan appropriately labelled, and how different the building is from the doctors' and the patients' spaces (the latter being the waiting room and the outside).

These differences arise largely from the very strong 'genotypical' spatial requirements that this type of interface commonly imposes. There are four main elements:

- the patients have to be held in an easily controlled space shallow in the building;
- the entrances to doctors' rooms have to be separated from the entrance to the patients' space by some kind of distancing device, usually an axial discontinuity of some kind;
- contact between doctors, and between them and other inhabitants, has to be possible without the possibility of accidental contact with patients;
- for the same reason, doctors have to have a separate entrance and independent routes to their part of the building – a common device for preserving status which we call the 'stage door effect'.

All of these requirements can be stated and realised in terms of the spatial elements and relations we have defined. Collectively they have a very powerful effect that makes the building very unlike the original house that was adapted for medical use. It becomes a building with a very *strong programme*, in the sense that virtually everything that occurs is specified and inscribed into the spatial pattern. Very little that is unprogrammed can occur. The strength of the genotypical model is such that

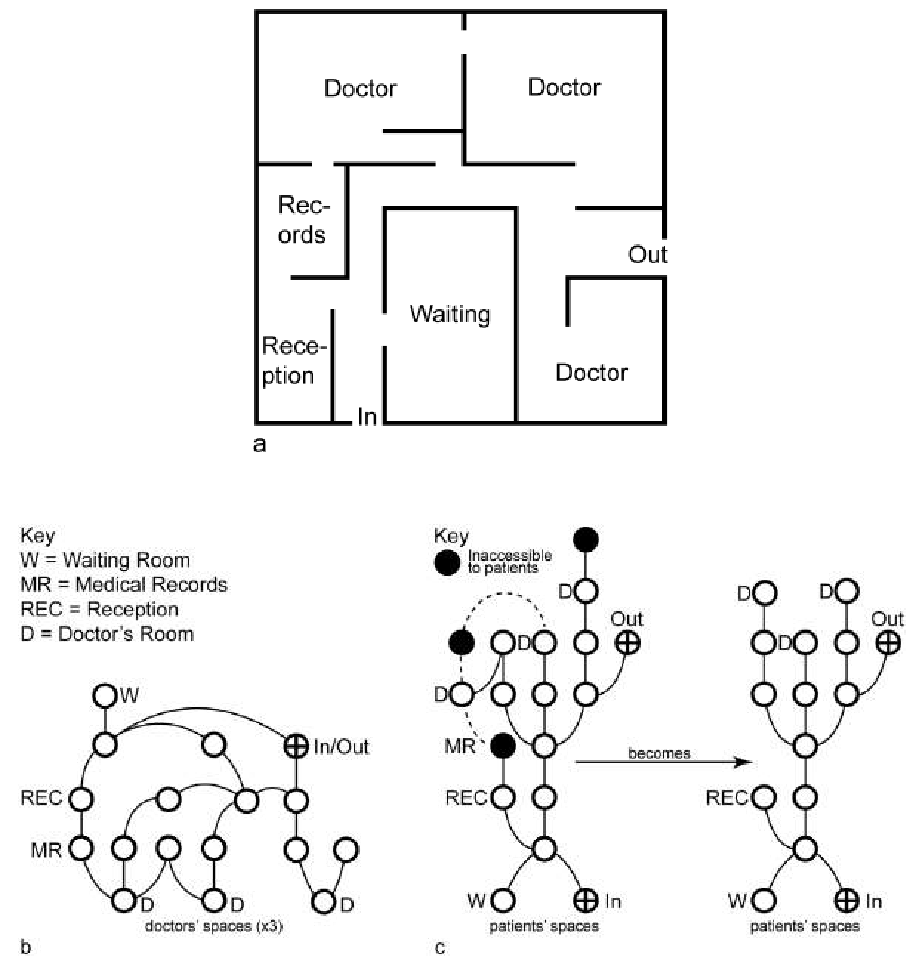


Figure 8.6 (a) Plan of doctors' surgery and (b) justified graph from the doctors' spaces, RRA = 0.475; (c) justified graph from patients' spaces, RRA = 1.693.

if one or other element is missing from the spatial organisation, then in all likelihood it will be compensated for by some behavioural practice or device.

The four elements of the model really come down to two general 'interface' requirements: the need to preserve the status and solidarity of the main 'inhabitants'; and the need to control the movement of patients. Insofar as the doctors' status is to be conserved by segregation and depth into the building, the patient must enter into the building; but, insofar as the patient is to be controlled, then the circulation from his entrance is controlled by the receptionist. The same circulation space thus works in opposite ways for doctor and patient. For the doctor, it is his means of distancing;

for the patient it is his means of control. For the doctor, the circulation is for the patient to travel to his segregated domain. For the patient, it is the means by which he is spatially restricted and confined to a room until his time comes. The relations of room to circulation are, as it were, reversed for patients and doctors. As we will see, under some conditions this reversal becomes a major theme in medical – as in other – interfaces, bringing about major discontinuities in the building form.

Figure 8.7 is a larger medical interface building selected, we admit, to illustrate a point. At first glance we might take it for just another building with a more extended circulation system appropriate for its size. In fact, a more careful examination will show that its global form exactly, and only, reproduces the four genotypical elements that we detected in the doctors' surgery. The geometry of the layout is quite different, but the 'syntactic' (all our spatial parameters) principles are identical. The only substantial effect of size would seem to be the duplication of the waiting area and the removal of its boundary. In this sense, the building increases the control of the patient and the 'reversal' effect in that, locally at least, the function of the circulation system is to create an area where patients can be both static and surveyed.

Figure 8.8 is another kind of larger medical interface building; the outpatients' department of a famous modern hospital, selected again, we admit, to illustrate a point. At first sight, it seems that we must give up our theory of the genotypical elements of the medical interface, since none at all seem to be present to any degree in this plan. Patients are spread through the building from shallow to deep; there is no distancing of doctors' rooms – on the contrary, they interface directly with the

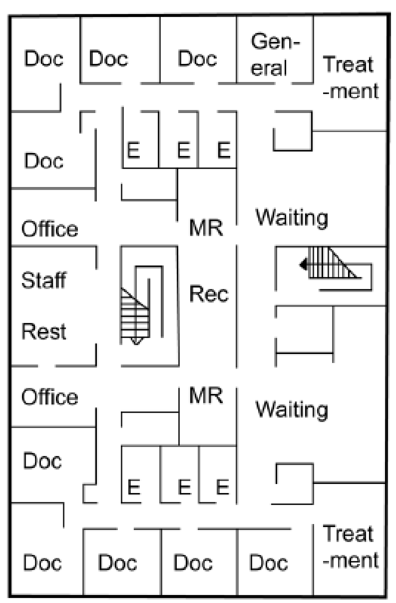


Figure 8.7 Plan of a 'category' type health centre.

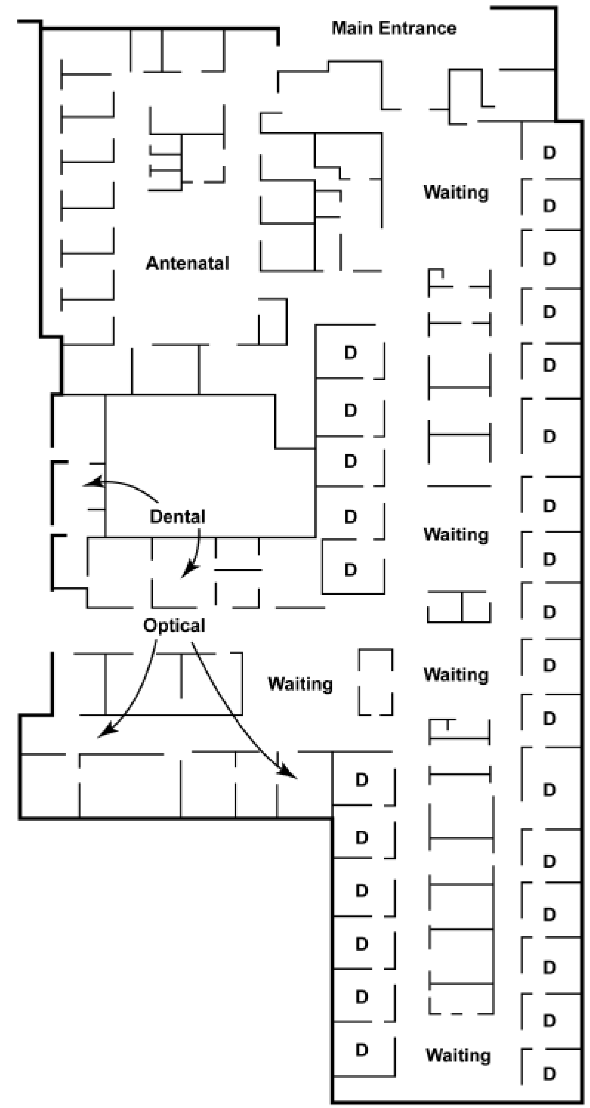


Figure 8.8 Hospital outpatients' department.

patient areas; there is no separate circulation system, apart from links from one room to its neighbour; and there are no stage door effects to speak of.

But a moment's thought about the social ideas underlying the genotype suggests another possibility. Surely what has happened in this case is that the genotypical elements that have to do with the distancing and reintegration of doctors as professionals have been dropped, and those that have to do with the control of patients have taken over. The increase in axial organisation, so that the plan becomes very shallow from the outside and most points inside, the direct interface with doctors' rooms, the splitting of patients into dispersed groups, could all be thought to have this effect. If, in the previous case, the global pattern of circulation was defined by the issue of professional status, in this case, it is dominated by the pragmatics of control. This is what we mean by a morphological discontinuity. The two types of solution cohere as part of a general theoretical model; but as actual spatial types they are utterly unlike each other.

Figure 8.9 is then a more recent solution to the same problem: the OPD [outpatients' department] of one of London's most recent hospitals. Once again we seem to have a morphological discontinuity. The original genotypical elements are not 'there' to any degree, but nor are the principles of the first OPD. In fact, the very contrary principles seem to be at work. The main effect of this rather labyrinthine spatial planning seem to be to segregate the patients, rather than the doctors, and to separate them from the outside as strongly as they are separated from each other. A greater contrast both in working, and in intuitive feeling, with the first OPD can hardly be imagined. Do these feelings have an objective basis? And does this have anything to do with the relation between spatial form and function in buildings?

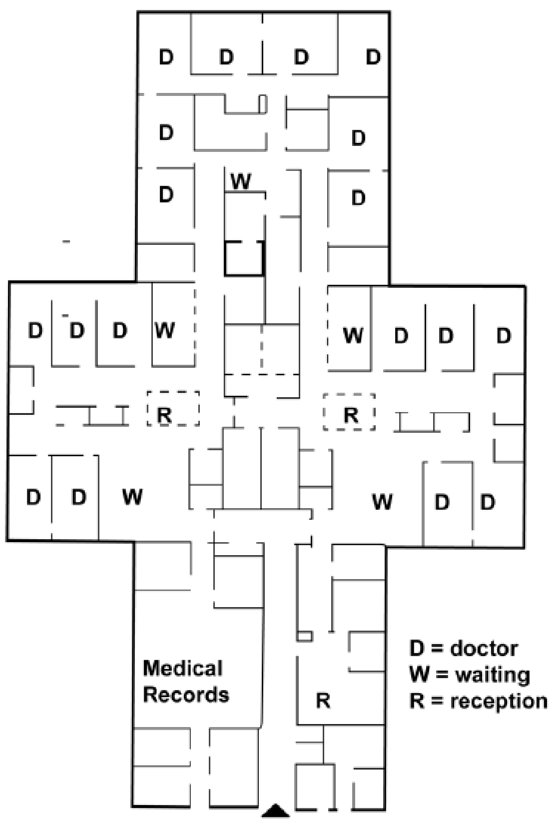


Figure 8.9 Hospital outpatients' department.

The answer to this question is difficult to formulate, but is nevertheless, we believe, vital to our understanding of large building complexes and why some don't seem to work as well as others. It runs something like this. As buildings grow larger, it becomes more and more difficult to maintain them as 'strong programme' buildings, that is buildings where most of what happens is specified by explicit or tacit rules, and built into the spatial structure of the building. There is a simple reason for this. As the numbers of people and the number of spaces necessary to accommodate them increase, so the amount of unprogrammed contact as the natural by-product of functionally defined movement is also likely to increase. The same effect is strong in settlements as they become larger, and indeed the liveliness that we often find in settlements is often largely a by-product of this process.

Our analyses of settlement form, together with our observations and simulations of them, have suggested, in fact, that settlements take the spatial form they do largely in order to generate, control and render predictable the unprogrammed by-products of complex patterns of spatial activity and movement. Settlements being on the whole 'weak programme' spatial complexes use the natural laws that relate spatial parameters to movement probabilities to create a kind of order in what might otherwise be an unmanageably complex and unintelligible social and spatial environment. The laws of space literally become the basis for a certain kind of informal order to take over where social programmes are weak. In settlements, all that is predictable is the general spatial distribution of people, and that is predictable, intuitively as well as formally, from the spatial pattern.

We believe that large buildings also have this character and this problem. Those that solve it by using the global structure of the building to generate and control a pattern of movement and potential encounter over and above the specific and localised functional programmes acquire a kind of spatialised culture by which we begin to recognise them as institutions. UCL itself is an example of such a complex, and so is the AA [Architectural Association]. Those that pursue a localised function at the expense of such a global patterning will not generate such a spatial culture and will be experienced as fragmentary institutions.

This may sound like a subjective evaluation, but we believe that our work on settlement forms and more recent work on buildings (especially an admirable study by a current MSc student of ours, Alan Penn[7]) can begin to provide an objective basis for such views. Take, for example, the axial maps of the two OPDs (see Figures 8.10 and 8.11). The integration analysis of the first shows an 'integration core' that links the outside to the inner recesses of the building and gives the plan a strong global structure. The second, on the other hand, has a core which is both deep in the plan and limited to its central areas. Bearing in mind that integration values are by far the most powerful determinants of global movement patterns, it is clear that all categories of user are going to be pressed towards relative immobility by the structure of space in the second case, whereas in the first they will be drawn

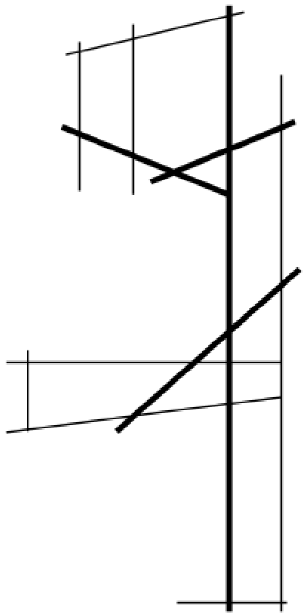


Figure 8.10 Axial map of Figure 8.8 with 'integrating core' marked by thick lines.

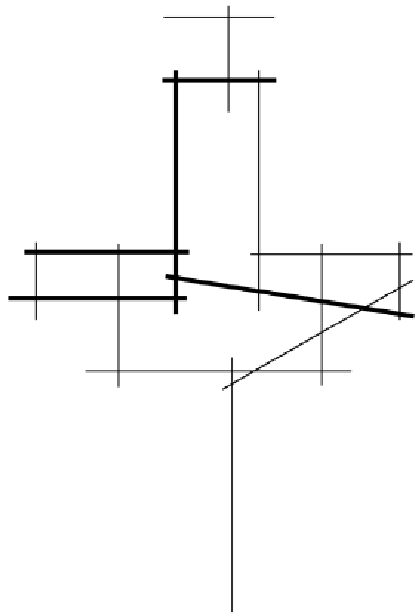


Figure 8.11 Axial map of Figure 8.9 with 'integrating core' marked by thick lines.

naturally into global patterns of movement and encounter. The second may be better functionally at the local level. But it achieves this at the expense of building the global institution. The first may be functionally more difficult to manage, but it will contribute much more to the global institution.

These suggestions are confirmed powerfully by an examination of the whole floor hospital plans of which the two OPDs are a part. In the first case, it can be seen that the interior of the OPD itself forms a significant part of the integrating core, a core which forms a strong grid reaching deep into the building. In the second case, the core has a simple tree form based on a very powerful single main 'street', and in this case the OPD – a major source of moving visitors to the building – is in a highly segregated zone, with no penetration at all from the core (see Figures 8.12 and 8.13). We have clues to believe that these powerful – and objective – differences in spatial structure, which are rendered visible by a theory of description, are powerfully implicated in the utterly different impressions that these two hospitals make both as functioning organisations and as spatial institutions.

What we are saying, in short, is something like this. In a small house, the role of space is not so much to reflect everyday functional patterns, but to add something to those patterns that would not otherwise be there. Because we have a best room, we

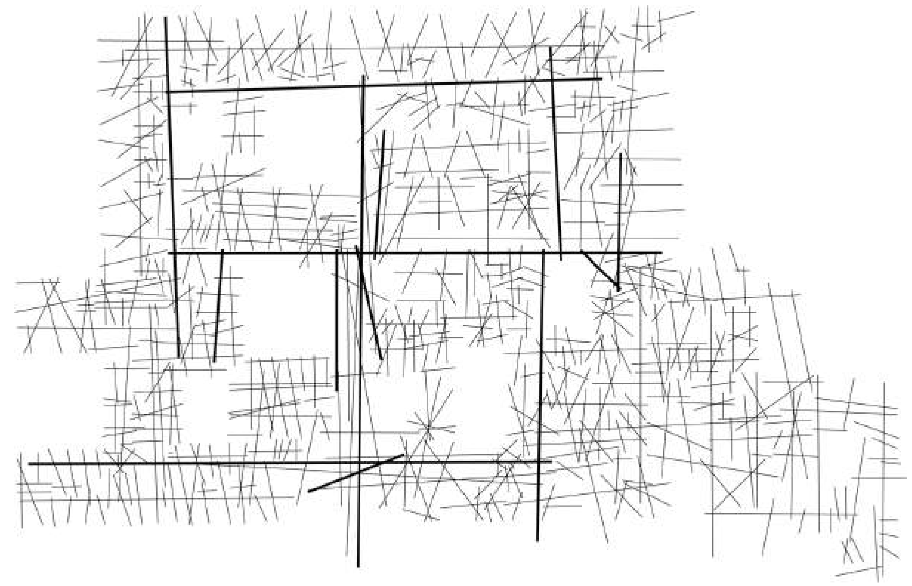


Figure 8.12 Axial map of whole hospital with 'integrating core' marked by thick lines.

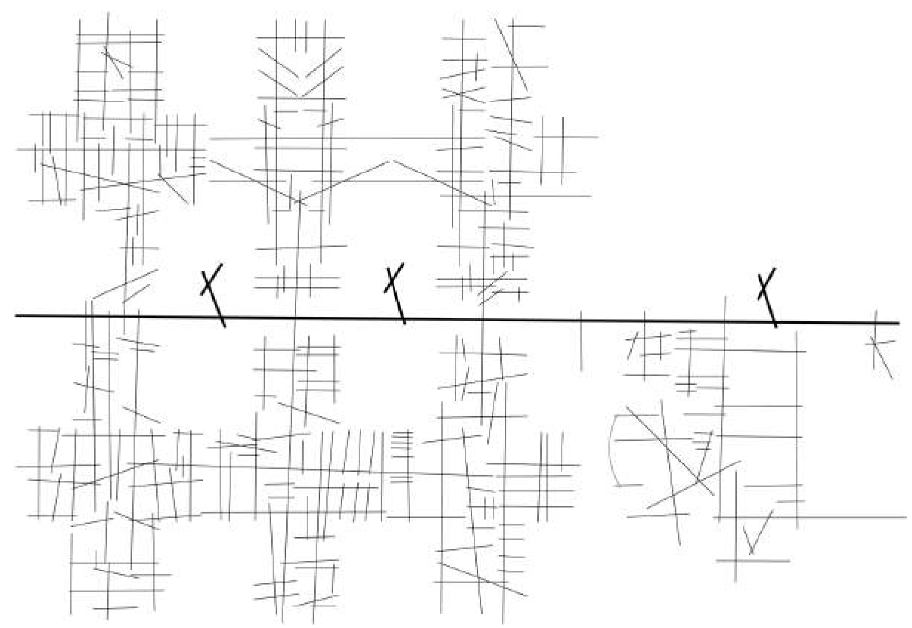


Figure 8.13 Axial map of whole hospital with 'integrating core' marked by thick lines.

do not need to go there, and we relax more. We have the space as a resource when we need it. Space *adds* a degree of extra culture to life, and life is the more pleasant because of it. At a rather larger scale, we find the 'strong programme' building, where everything is specified, and where typological-functional analysis is possible on the basis of interfaces, morphological discontinuities and social strategies. But with the large building we find a new phenomenon. Again, space is adding something to function, as with the house, but here it is not adding the structure of a given culture, but something like the opposite. It is adding the generation of a social field that is unstructured, but which, like with a settlement, acquires a predictability and a reproducibility – and therefore a social identity – through adapting its spatial organisation to nurture and organise this emergent phenomenon. The true function of large building complexes in our time is, we believe, to create these emergent social organisms.

Notes

1. B. Hillier, P. Stansall, and J. Hanson, *The analysis of complex buildings* (London: Unit for Architectural Studies, Bartlett School of Architecture and Planning, University College London, 1978).
2. A. Forty, *Words and buildings: A vocabulary of modern architecture* (London: Thames and Hudson, 2000).
3. P. Frankl, *Principles of architectural history: The four phases of architectural style, 1420–1900*, translated by J. F. O'Gorman (Cambridge, MA, and London: MIT Press, 1973) (German original 1914); C. Rowe, *The mathematics of the ideal villa and other essays* (Cambridge, MA: MIT Press, 1976).
4. B. Hillier and J. Hanson, *The social logic of space* (Cambridge: Cambridge University Press, 1984). [At the time of publication of this piece, this was 'forthcoming, 1983'.]
5. B. Hillier and J. Hanson, 'Discovering housing genotypes', Working paper of the Unit for Architectural Studies, University College London (1982). (The editors suggest that readers may wish to consult Hanson's *Decoding homes and houses* (Cambridge: Cambridge University Press, 1998) for further elaboration of these ideas.)
6. R. B. Wood-Jones, *Traditional domestic architecture of the Banbury region* (Manchester: Manchester University Press, 1963).
7. Editors' note: see contribution by Professor Alan Penn in the introduction to Chapter 2 of this volume, 'The man–environment paradigm and its paradoxes', and as co-author to the articles reproduced in Chapters 14 and 16.

9 Quite unlike the pleasures of scratching (1985)

Introduction to 'Quite unlike the pleasures of scratching: Theory and meaning in architectural form'

Wilfried Wang

The development of space syntax theory and methodology at the then School of Environmental Studies, later the Bartlett School, University College London, in the second half of the twentieth century under the intellectual guidance of Bill Hillier and Julienne Hanson set the frame for fundamental research into the social significance of the built environment. Bill Hillier's ambitions are expressed in the title of the essay: in the 1980s Hillier wanted to enter the more orthodox discourse on the making and meaning of architecture.

The 1980s were the years in which academicised Modernism was being challenged by realised examples of Postmodernism. Two decades earlier, sceptics of Modernism had succeeded in finding a wider audience for their critique of technocratic and mechanistic rationalism. Where the interwar architects such as Hugo Häring, Josef Frank or Alvar Aalto had failed to rally support for their alternative conception of modern architecture, the two-pronged attack on Modernist architecture and urban design by Robert Venturi with his *Complexity and contradiction in architecture* and Aldo Rossi with his *L'architettura della città* (both published in 1966)[1] set off convulsions in architectural and urban design theory and ultimately practice.

Bill Hillier's essay claims a place for the space syntax approach in the regaining of lost meaning in architectural form against, on the one hand, the anti-Modernist position of Roger Scruton, as laid out in his book *The aesthetics of architecture* – essentially the mouthpiece of the then Prince Charles, which seeks to reestablish classical architecture as the norm – and on the other hand Alberto Pérez-Gómez's attempt to breathe life into depleted architectural meaning by restoring the role of phenomenological perception as the prerequisite to the poetic-metaphorical construction of meaning in architecture.[2]

Space syntax at the time had already proven its ability to identify the structure of socio-cultural significance as embedded in the built environment. Its theory and analytical methodology could be used to read these meanings in any building as well as in any urban fabric. For example, in the case of concert halls it could be shown that the disposition of the foyer, auditorium and backstage areas related to the relative roles of the audience and musicians at the time of the buildings' conception. Analysed together with the way the concert halls were used, the mapping of social values into the respective buildings could be retrieved. The Royal Albert Hall thus is the pure representation of segregated Victorian class structure, while the Philharmonic is the ideal representation of the birth of a democracy.

Hillier's ambition at the time, as documented with this essay, was to be able to lay out a theory of a social logic of form. For this purpose, he not only used Scruton and Pérez-Gómez then widely read interventions as a point of engagement in orthodox architectural theory, but, especially in Pérez-Gómez's case, probed beyond the main writing to demystify the premises upon which they were based. In the latter case, therefore, Pérez-Gómez's principal touchstone – the theory of phenomenology as proposed by Edmund Husserl – was subject to Hillier's incisive analysis. Hillier holds that it is impossible to suspend the conceptual frameworks that observers possess in order to reach an understanding of the 'pure essences'.

The fact that Hillier concerned himself in this essay with both Scruton and Pérez-Gómez's focus on mathematical and geometric definitions of architectural form is unfortunate, since by that time the maturation of space syntax had permitted the methodology to overcome both the mathematical and geometric specificities of any given space, while uncovering the space's essential socio-cultural meaning. At the same time, Hillier realised that space syntax methodology was reduced to the two-dimensional, planimetric level; it was anything but spatial. In part, this realisation propelled Hillier to seek ways to speak analytically about the meaning of architectural form.

In discussing architectural theories, it is possible to understand the majority of treatises, especially the European classics from Vitruvius through Alberti to Schinkel, as offering an idea of how architecture is possible and – not only in the classical texts – giving key components with which architecture could and, in some cases even, should be made. In other words, architectural theories were typically concerned with providing a descriptive range of the task with a few prescriptions. In Le Corbusier's case, for example, the prescriptive five points of architecture replaced the prescriptive classical orders. While with the classical orders there was a shorthand for rudimentary dimensions of meaning, such as the Doric being sturdy and masculine, and so on, the Corbusian five points of architecture did not claim any such engendered meanings.

No architectural theory has been able to explain the meaning of any given design; no architectural theory has offered an analytical methodology to decipher the significance of any given architectural composition. Neither Scruton nor Pérez-Gómez set out a descriptive methodology in their books under discussion, and neither did Hillier. To be fair to Hillier, however, he did outline his vision of an analytical methodology: 'intelligibility, genotypical forms and genetic construction' would form the basis of Hillier's approach to describe architectural meaning. In this approach, he mirrored the theoretical underpinnings of space syntax.

The key problem with this approach lay in Hillier's reluctance to address the description of architectural forms as such. Instead of this task, Hillier considered the notion of 'genetic construction' sufficient to provide a *logical* description of form. However, such a logical description is insufficient to offer a precise understanding of any given architectural composition.

Whereas architectural design is synthetic, that is, architectural design composes material and formal components, architectural theory's task is to provide analytical tools. While architectural design seeks to create socio-cultural meaning, architectural theory should be able to dispassionately explain how that meaning is constituted through any given composition. The practice of design is the reverse of the practice of analysis. Neither requires an explicit theory; however, both architectural analysis and architectural criticism should be based on an explicit descriptive method, wherein every single component of a design can be understood in its relation to the other parts and to the overall whole.

Bill Hillier was a thinker in the realm of environmental design. He sought to marry his theory of space syntax with an equivalent theory of architectural form. That is the legacy of this essay, and it remains a task for others to complete.

Quite unlike the pleasures of scratching

Theory and meaning in architectural form

Bill Hillier

Originally published: Hillier, B. 1985. 'Quite unlike the pleasures of scratching: Theory and meaning in architectural form'. *9H*, 7, 66–72.

Recent critics of Modernism have either mystified or denied the central issue of architectural theory: the analysis of order in architectural form. This has happened because there is a crisis in architectural theory itself, brought about by the divorce between theories of form and theories of meaning. This divorce has resulted from the mathematical bias of theories of form and the linguistic bias of theories of meaning.

This article reviews the arguments of two recent architectural critics, Roger Scruton and Alberto Pérez-Gómez, and their respective philosophical sources, in an attempt to clarify – or, more appropriately, demystify – their violently anti-theoretical positions. By reviewing the writings of Plato, Vitruvius and Alberti – and showing how they have often been misinterpreted and misunderstood – the author suggests that architectural theory has always preoccupied itself with a rational description of form to show how architecture can be both intelligible and meaningful to us.

It is argued that the study of form and meaning can only be unified by a syntactic approach to the understanding of form itself, and only then can we begin to address the problems of form raised, but not solved, by Modernism.

Theories and meanings

Theories are bundles of abstract, interrelated propositions that we use to give order to the evidence of our senses. They provide, as it were, an abstract picture of what the world is like, which we then use to interpret our concrete experience of it. Theories are closely bound up with the idea of *meaning*. What the world means to us is, at least in part, a product of the theoretical constructs that we bring to bear on it.

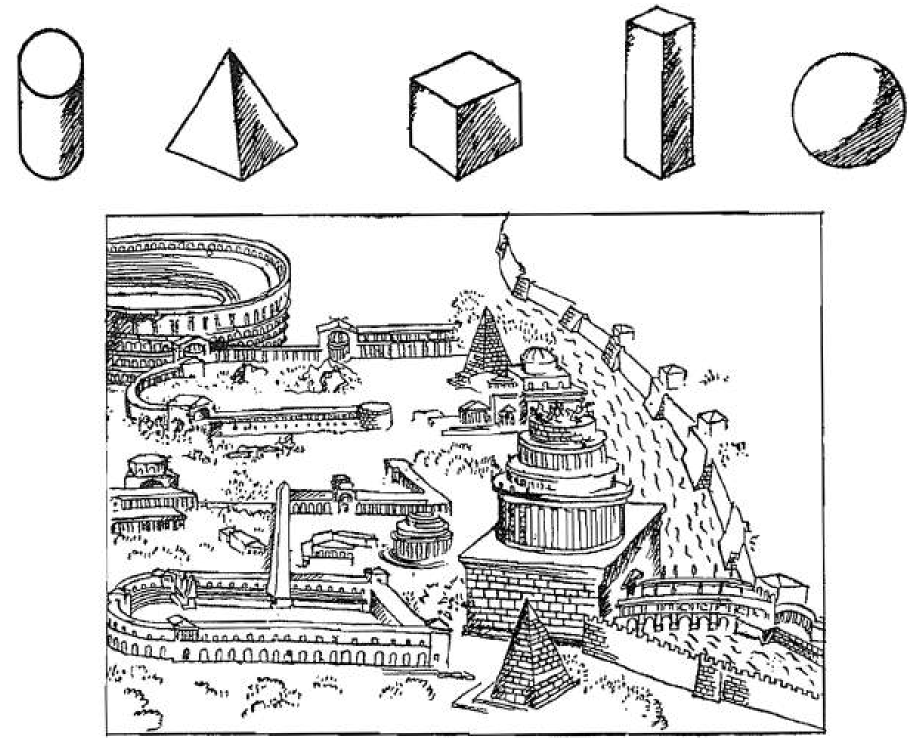


Figure 9.1 Le Corbusier: 'The lesson of Rome'.[3]

Now, if twentieth-century philosophy can be distinguished from previous times by anything at all, then I would suggest that it lies in the recognition that this link between theory and meaning applies to everyday life and common sense as much as to science and philosophy. The sense that we understand the world – whether it is the physical world or the social and cultural milieu in which we live – depends on the theory-like constructs that we impose on our experience of it. Without theory, the world and everything in it, man-made as well as natural, becomes meaningless and unintelligible.

Two radical consequences follow from this insight, each providing a characteristic theme of twentieth-century thought. The first is that science is seen as not unlike everything else that people think and do, but as having a natural continuity with it. As an activity, science is no more than a special extension of common sense, in which we try to test our theories against experience, and where these theories are found wanting, to conjecture new ones and then test those. The epistemological claims of science, thus, rest not on some concept of *truth* but simply on more developed and better tested theories.

The second consequence is that if everyday life and culture carry social meaning for us through theory-like constructs, then it ought to be possible to investigate how this is so: how it is that everyday activities by which we 'read' culture – eating, building, dressing, interacting and so on – carry shared meaning for us. Or, put more precisely, what it is about the *form* of these activities that makes them *intelligible* as parts of culture. The name for this intellectual project is 'structuralism'. It asks: can we give a *description* of these activities in such a way – necessarily an abstract way – as to show how it is that we can find pattern and meaning in our experience of them.

Architectural theory: The central theme

Now, any architect who had read his Vitruvius or his Alberti might well feel that he had reason to feel smug at this point, because architectural theory has always been 'structuralist' in this sense. The central aim of architectural theory – from Vitruvius to Le Corbusier – has always been to arrive at some description of the abstract principles underlying *order* in architecture, with the assumption that, once understood, these principles would also provide us with an understanding of how it is that people derive cultural meanings from architecture, and how they find it intelligible in the first place. In this sense, architectural theory has always attempted to link the everyday understanding of buildings with the problem of giving a formal description of them. Concepts like proportion, symmetry, composition and so on all have this double intention: of characterising something about objective order present in architecture, and of providing insight to the theoretical preconceptions that people must have in their heads if they are to read such objective properties as constituting a form of meaningful order.

It has usually been on this basis that architectural theory has tried to establish systems of precepts for 'good' architecture. In this sense, there is quite a remarkable continuity from Vitruvius through Alberti and Durand to Le Corbusier. With all their differences in style of thought and social context, the *Leçons* and the *Modulor* are a continuation of the classical project of finding and clarifying the intellectual bases on which formal order can be linked to intelligibility and meaning, and then using this as a basis for a prescriptive system of design.

But the thematic continuity in architectural theory is not confined solely to the prescriptive project. There are also common analytic foundations. It has always seemed self-evident that if there is an underlying order to be found in the morphology of architectural form, then it must always be a *mathematical* order. Two possibilities have always co-existed, side by side. The first is based on numbers, and rests on the simple proposition that order in architecture arises from regularities in the ways in which we can combine numerical ratios with each other. Theories of proportion and modularity belong to this category. Ultimately, they are based on the 'Pythagorean' idea that order in *nature* – and especially in musical sounds – is founded on regular ratios that can be expressed as relations between whole numbers. In this way, numerical theories of order are proto-scientific: they postulate an underlying numerical order in nature, and propose to reproduce it in the man-made world of architecture.

The second possibility is based on geometry, and rests on similar epistemological foundations, though in this case the ultimate authority is Plato's theory of the geometrical nature of the 'deep structures' of the physical world. Again, its central proposition is that nature exhibits an underlying *geometrical* order, and it is this that we should seek to emulate in architecture. Once again, the theory is quite explicitly proto-scientific in its postulation of an underlying, describable, mathematical order which ought to be imitated in architecture.

Two rather different kinds of theory of architectural order arise from these mathematical possibilities. On the one hand, we have theories of how good architecture may be *generated* by a step-by-step process, following given rules. Modular and proportional theories are essentially of this kind: we might call them theories of the 'genericity' of order, noting that they say little about the final form of the building, but deal with the process of its formation. On the other hand, we have theories of what the *final form* of the building should be like – a state description, perhaps, rather than a process description of how we should arrive at the form – which tend to be of a geometrical rather than a numerical nature.

These two types of theory do interact. For example, the golden section founds a geometrically based theory of proportion, and therefore of genericity. Indeed, the almost mystical status of the golden section may be to do with the fact that it alone proposes a theory of order that is at once a theory of genericity *and* a theory of geometrical form.

Both types of theory, however, tend to lead to a single profound difficulty, one which may have prevented either theory from making significant progress over the past two thousand years. It is this. If the underlying order in the physical world is mathematical, and the aim of architecture is to reproduce it in buildings, it then follows that the order we are trying to impose on buildings is a *natural* order, rather than a social order. Insofar as we try to identify the abstract principles of architecture through mathematics, we are by implication *naturalising* – and thus *desocialising* – the notion of order in architecture. We find ourselves required to believe in eternal verities of form, rather than cultural and social relativities.

In at least one sense this flies in the face of common sense. It leads us to expect that our reading of architectural form will be indifferent to cultural meaning, but this is not the case. The degree and type of formal order a building possesses is one of the first clues we have to its social and cultural nature. We read buildings not as objects which remind us of the order of nature but as objects that bring forcibly to mind the order of culture. It seems unlikely, then, that mathematical theories can be telling the whole story, since they lead to theories that *remove* architecture from our *cultural* reading of it. In effect, they divorce the theory of order from the theory of meaning.

There is, unfortunately, an escape from this fundamental difficulty, one that has been often taken in the history of architectural theory. It is to argue that cultural meaning in architecture resides not in the form itself, but in the way in which that form is elaborated and decorated by *figures* of various kinds (to borrow Alan Colquhoun's useful expression[4]): that is, by characteristic ways of handling

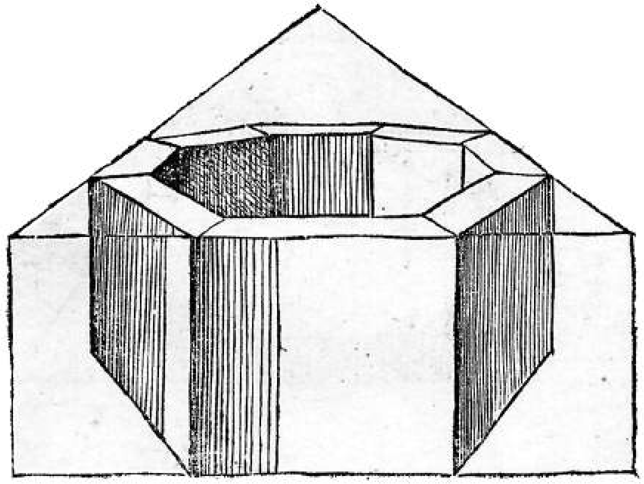


Figure 9.2 Octagonal ring in perspective (from Serlio 1566).

and relating constructional elements, openings and so on. Figures are so clearly influenced by historical periods as to leave their 'cultural relativity' in no doubt. They can, therefore, act in a way comparable to rhetorical 'figures' in creating a sense that social meaning is being transmitted. The classical orders are themselves the prototypes of this concept of meaning.

A direct consequence of this apparent solution to the difficulty is that the notion of meaning in architecture is historicised and made dependent on the contingencies of historical developments, such as constructional techniques or modes of production. It is this, I believe, that has rendered the study of architectural meaning so intractable. It has led architecture to fall prey, again and again, to the linguistic analogy[5] in which the false hope is set up that architecture will recover meaning by first identifying the word-like units that make up architecture, then establishing what they mean, and then showing how they may be combined into sentences to convey cultural meanings. Why the linguistic analogy is false to architecture – and why it has always failed to deliver even a fraction of its promise – requires us to penetrate a little more deeply into what we mean by *meaning*.

Significance and signification

The notion of meaning, in architecture as elsewhere, raises two quite distinct possibilities which we might sum up as the difference between *significance* and *signification*. Theories of *significance* are those where architecture is held to refer first and foremost to itself, to its intrinsic nature as architecture, as an instance of the field of ordering possibilities that we call architecture. Theories of *signification* are those where architecture is held to stand for something other than itself: power, religious ideals, social statuses and so on. We might say, more succinctly, that significance is about *syntax*, that is, about shape, form and pattern in architecture as a thing in itself; whereas signification is about *semantics*, that is, about the extrinsic domains of meaning that are awakened as a result of that syntax.

We may also say that in architectural theory – as we have inherited it – the question of significance is dealt with mainly through the mathematical preoccupation with *form*, while the question of signification is dealt with primarily through the linguistic preoccupation with *figure*. This is as disastrous as it is illogical. The field of architectural theory – focusing as it must on the analysis of form and meaning – is, in effect, split into two irreconcilable camps; neither of which can make contact with the other, and neither of which can progress without the other. Common sense would suggest that a theory of the social signification of architecture can only be based on a theory of its significance in itself in the first place. We must, if you like, have a theory of how architecture can mean anything at all before we can have a theory of what architecture might actually mean. Yet the epistemological split between the two camps precludes any such possibility.

Scruton's critique of architectural theory

The theoretical impasse that architecture has engineered for itself has had the effect of rendering it highly susceptible to external attack. Roger Scruton, for example, in his *The aesthetics of architecture*,[6] can, as a philosopher, mount an attack which is not only an attack on Modernism in architecture, but also an attack on the whole idea that architects can, by the rational analysis of architectural objects, arrive at a better understanding of what order in architecture comprises, and how, therefore, it has meaning for us. It is, in other words, an attack on the possibility of even having a theory of architecture.

Scruton's attack is quite comprehensive, and he exploits the very weakness of architectural theory that I have described. For example, he attacks all architectural theorists who have tried to objectivise the understanding of form, through concepts such as proportion, symmetry and so on. The meanings of these terms, he argues, insofar as they apply to buildings at all, do not refer to the *form* of the building, but to its detail. Architectural theory – which he defines contemptuously as 'usually the gesture of a practical man unused to words'[7] – has been mistaken in trying to find meaning through the abstract, rational analysis of architectural morphology. 'Appropriate' detail is all that is involved, since architecture is neither art nor science but simply an expression of polite and decorous behaviour, like laying tables or knowing about wine.

But Scruton wants to go even further than saying that the analysis of built form is useless; he wants to show that it is also *philosophically* misconceived. Rational enquiry into architecture, he suggests, should be aimed, not at architectural objects, but at the nature of our experience of those objects, since: '. . . concepts like the appropriate and the beautiful . . . take their sense not from the objects to which they are applied, but from the state of mind which they serve to articulate'[8] and 'the concept of the "appropriate", like other concepts which articulate aesthetic response, eludes explicit definition'.[9]

In arguing this, Scruton is, of course, repeating a major principle of Kantian aesthetics, which properly argues that terms we use in order to express *judgements* about objects should not be thought of as properties of those objects themselves, but as aspects of the way in which we think about them.

Now, Scruton wants us to extend this principle to concepts like 'proportion'; that is, to treat it as a word which describes our experience of objects themselves. This is not because he regards the concept as unimportant. Far from it: '. . . it is precisely because proportion is so aesthetically fundamental that we should beware of tying it down to an explicit definition.[10] But '. . . there is never any need, in the practice of aesthetic judgement, to say what "proportion" really means in the sense of pinning down the term to some one property or set of properties which provide its true aesthetic sense'.[11]

It is here that the limitations of Scruton's ingenious attack suddenly show themselves. Of course, there is no need to define terms like 'proportion' in the practice of *aesthetic judgement*. But what about the practice of *design*? Can the designing architect – who, after all, is responsible for giving Scruton his experience – be satisfied with purposely undefined terms that articulate his own feelings more than they articulate the properties of the buildings he is designing? The contrary is surely the case. The bane of architectural theory is that it can turn the intuitive insight of the mind experiencing architecture into a reflective and descriptive comment on architecture itself. In design, the striving for rational understanding does not result – as Scruton seems to think – from a naïve and overweening desire to reduce the world to a codified system, but from the simple need to be able to create new experiences.

Scruton eventually shows up the paradox in his argument himself: 'What is meant by proportion can be understood as we have understood the concept of the appropriate. Proportion is exhibited by a building whose parts – judged in terms of their shape and size rather than in the terms of their ornamentation – provide visual reason for one another.'[12]

'Provide visual reason for one another'? Surely, this is a trick of words if we are meant to infer that such provision is only an aspect of judgement, and not in the least an aspect of buildings. If it is an aspect of buildings, then the entire problematic of architectural theory rises before us, since the expression can only refer to some abstract configurational property of the form itself and cannot, by Scruton's own account, be reduced to a matter of 'appropriate' detail. The whole aim of architectural theory has, one might say, always been to try to find out what this type of expression means in terms of designable buildings.

In the final analysis, then, Scruton's philosophy is not only useless to the architect but also bizarre. He opposes the rational analysis of the architectural object on no better grounds, so far as I can see, than that he is a Kantian philosopher

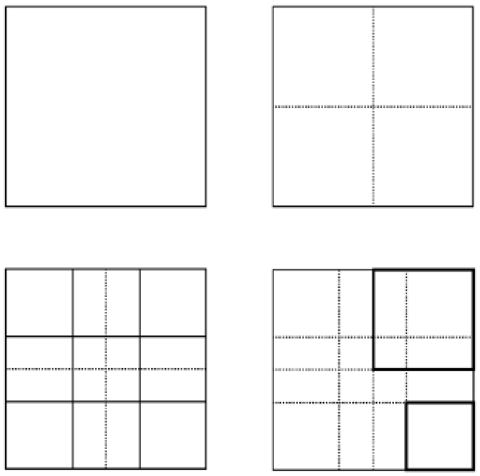


Figure 9.3 Schematic analysis of geometrical construction of Malevich's composition of 1915: red square and black square: *Painterly Realism of a Boy with a Knapsack – Colour Masses in the Fourth Dimension*, MOMA.

of mind. More strangely still, he tries to recruit Alberti – on whom Scruton is an acknowledged expert – to his obscurantism. I have always found Alberti particularly clear in opposing the points that Scruton would need him to make. 'The parts', Alberti writes in a well-known passage, 'of which architectural forms consist are lines, angles, extensions, and the like'.[13] For him, these abstractions must be considered first 'for it is undeniable that there may be, in the mere form or figure (meaning overall configuration) of a building an innate excellence or beauty, which strikes and delights the mind where it is perceived as much as it is missed where it is not'. And, as though reproving Scruton in advance for subordinating form to detail, he goes on: 'We should erect our building naked and let it be quite completed before we begin to dress it with ornaments.'[14] And, as though to confirm the modernity of his structuralism, he adds: 'But the judgement which you make that a thing is beautiful, does not proceed from mere opinion, but from a secret Argument and Discourse implanted in the Mind itself.'[15]

The phenomenological critique

An even more extreme attack on the rational analysis of architectural form has also come recently from a very different philosophical quarter, phenomenology, with the publication of a book by Alberto Pérez-Gómez titled, after Husserl, *Architecture and the crisis of modern science*.[16] Actually, the book is not about

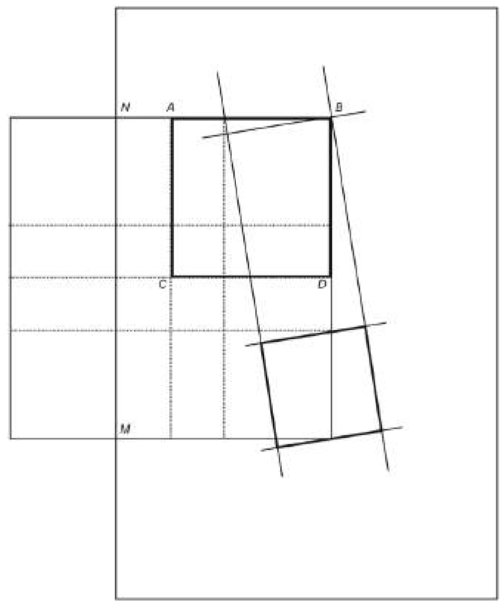


Figure 9.4 Schematic analysis of geometrical construction of Malevich's composition of 1915: red square and black square: *Painterly Realism of a Boy with a Knapsack – Colour Masses in the Fourth Dimension*, MOMA.

science at all, but about geometry in architecture, and its relation to meaning, or, more precisely, the loss of meaning. Pérez-Gómez holds a kind of 'lost innocence' theory of architectural meaning, with science as the leading ravisher. According to Pérez-Gómez, until the advent of science '. . . Architectural intentionality was transcendental, necessarily symbolic',[17] and 'Geometry and number, prototypes of the ideal, since time immemorial have been symbols of the highest order.'[18] But:

> '. . . the "new science" of Galileo . . . implied a radical subversion of the traditional astrobiological world view . . . and . . . pretended to substitute for the reality of the live world, infinitely diverse, always in motion, and defined essentially by qualities, a perfectly intelligible world, determined exclusively by its geometrical and quantitative properties'.[19]

Gradually, with Durand playing the role of Mephistopheles, the symbolic geometry of the 'mythical ancient world embodied in the writings of Vitruvius'[20] came to be debased to a mere 'instrument for the control of technical operations'[21]; while '. . . things became numbers, not understood as their Platonic or Pythagorean transcendental essences, but as objective and intelligible forms'[22] and, in the final analysis, it was all due to the fact that '. . . the new science of Galileo postulated the initial split between the perceptual and the conceptual spheres of knowledge'.[23]

On the basis of this analysis, Pérez-Gómez condemns all rational and abstract analytic thought about architectural form, whether 'typological or morphological'[24] as leading away from an understanding of meaning. The argument is one that will appeal immediately to many. The idea that the prosaic, materialistic world of today has somehow forced out of architecture a direct experience of deep symbolic realities will strike a chord, and not only among those who feel intellectually dispossessed by science. Even more appealing is the promise that the revival of some archaic (and arcane) geometric symbolism might save architecture from Modernist functional shapelessness.

The trouble is that none of it is true. The world of symbolic geometry cannot be found in the writings of Vitruvius. Nor can the rake's progress of rational thought that Pérez-Gómez describes be found in the history of philosophy and science. As with Scruton, Pérez-Gómez's ideas (as his title suggests) are derived more from philosophical principle than analysis of evidence. To understand this particular line of attack on rationality we must understand the strange – and strangely persistent – early twentieth-century philosophy of *phenomenology*.

Phenomenology, in the sense that Pérez-Gómez means it,[25] was developed by Edmund Husserl in the early part of the twentieth century. Husserl was nothing if not ambitious. Like other philosophers he thought that it was the job of philosophy to attend, not only to what we think of, but also to the mental and conceptual apparatus we think *with*. Like others, he saw that the 'evidence of our senses' was a product of the theoretical frameworks we use to obtain it. But, unlike others, he thought that the problem could be overcome, not by analysing these frameworks, but by *by-passing* them.

This was to be done by what he calls 'bracketing' or 'suspending the natural attitude', natural attitude meaning the simple-minded way in which we take our everyday perceptions and concepts for granted. Bracketing can be explained quite simply. If I write 'I see you', then we are likely to take for granted what I mean. But if I write 'I "see" you', then to whatever was previously meant is added (through the inverted commas) that I call into question what the word 'see' means. By applying this trick to everything, Husserl thought that he could arrive at 'pure essences' or 'pure phenomena', which were free of presuppositions and independent of conceptual frameworks. Then, on the basis of these 'pure phenomena' he would erect an *a priori* (that is, prior to experience) *science* of cognition, which would act as a kind of master science and compel the rewriting of all the existing sciences, such as physics and biology. To use his own words:

> 'the term "phenomenology" designates two things: a new kind of descriptive method which made a breakthrough in philosophy at the turn of the century (a somewhat immodest claim, since he invented it himself!), and an a priori science derived from it; a science which is intended to supply the basic instrument for a rigorously scientific philosophy and, in its consequent applications, to make possible a methodological reform of all the sciences'.[26]

Husserl thus combines an almost Frankenstein-like belief in the possibilities of his super-science with a fundamental attack on all existing sciences for their allegedly naïve and 'positivistic' view of phenomena and rationality. It is, of course, only the latter anti-scientific stance that writers like Pérez-Gómez take up in order to support their attack on rationality. He points to the existence of a philosophy which is supposed to have demonstrated the inadequacies of science, and argues that this is because science and *meaning* are incompatible with each other. It would obviously be very inconvenient to such an argument if the founder of the philosophy were in fact arguing that we should replace the existing sciences with a single super-science of his own. This suppression of the 'other half' of Husserl does, however, have the useful effect of making the philosophy look less crazy than it actually is.

The whole argument would collapse if science were not 'positivistic' in the sense of being naïve about the relation between thought and reality. And, of course, it is not. The whole aim and justification of science has always been to call into question our existing theoretical constructs about the world, and, where possible, to try to conjecture and test new and better ones. The starting point of science, therefore, is the same as Husserl's: the realisation that our theoretical presuppositions are not reliable guides to what the world is actually like. Where Husserl differs from science is that he thinks some kind of ultimate *truth* is possible, whereas science continues to rest on theories for its 'piles in the swamp'.[27] Phenomenology is, probably, the last refuge of the old idea that there might be a scientific certainty to replace old certainties of religion – and presumably this is why the philosophy is so attractive to those with a mystical turn of mind (including many on the extreme right), where the uncertainties of mere theory are unacceptable. Whether or not this is true, Pérez-Gómez's argument is phenomenological not only in that it criticises scientific rationality but also in that it proposes some ultimate architectural truth, while remaining appropriately vague as to where this truth might be found.

But even more damaging to Pérez-Gómez's case against rationality is that this 'mythical ancient world' of pure geometrical symbols seems to be entirely imaginary. One can search the *Ten Books* of Vitruvius[28] from cover to cover (in which the 'mythical ancient world' is said to be found) and discover not one single reference to the *symbolic* importance of geometry, nor, indeed, to the symbolic importance of anything. There are, in fact, only three references to geometry in Vitruvius (his concept of proportion being entirely of a modular or numerical nature). The first reference is where he advocates the use of 'brazen gnomons' to decide the layout of streets in a new town, but not with any hint of symbolic intent, but merely to optimise (and I use this contentious term after much consideration) the 'shutting out of winds'[29]. The second is where he describes how both a square and a circle can be derived from the outstretched human body – as depicted in his well-known diagrams – but from this he derives not an argument for geometry, symbolic or otherwise, but an argument for proportion.[30] In his third reference, he does actually refer to Plato and Pythagoras, first to 'the very many useful theorems of Plato', which he saw as useful in solving dimensional problems of layout; and then to the theorem of Pythagoras, which 'affords a useful means of measuring many things, and is particularly useful in the design of staircases in buildings, so that the steps may be at the proper levels'.[31]

In other words, not only does Vitruvius not use geometry and number symbolically but, on the contrary, he does with them exactly what the moderns are accused of doing: he uses them as an 'instrument of control of technical operations'.[32] The point is hardly worth labouring. Vitruvius is a rabid functionalist, and his *Ten Books* read more like the back pages of the *Architects' Journal* than the poetic embodiment of a lost mystical world. Even the personality of Vitruvius, projected through the short introductory vignettes he writes for each book, is thoroughly modern. He makes it quite clear that one of his main aims in writing a learned treatise is to show how deserving he is of more work than has so far come his way.[33] At the centre of the supposed 'mythical ancient world', there is Vitruvius, touting for work.

Plato and intelligibility

It is hard to see how Pérez-Gómez could have misread Vitruvius so badly. But it is easier to forgive him for his misreading of what we might call the philosophy of geometry, since many myths on these matters already exist and dominate architectural discourse. First among these myths must be the strange – yet persistent – misunderstanding of 'Platonic solids'. The 'Platonic' or 'regular' solids – so-called because all their faces are congruent – are the tetrahedron, the octahedron, the hexahedron, the dodecahedron and the icosahedron, forms with relatively little visual interest and certainly of little direct use in the construction of architectural forms. They are known as the 'Platonic' solids not because Plato discovered the fact that only five regular solids were possible (the proof is in Euclid, but its origins are unknown) but because Plato used them in his *Timaeus*, in relation to the theory of the 'four elements' of earth, air, fire and water, to construct a theory of the micro-structure of matter. The theory was very simple, and extremely physical. Since all four elements behave differently, he argued, they must be constituted of different building blocks. Since there are only a few 'regular solids' then it is likely that different types of matter are built from different solids. 'Let us assign the cube to the earth', he says,

> 'for it is the most immobile of the four bodies, and the most retentive of shape, and these are characteristics which must belong to the figure with the most stable faces . . . while similarly we assign the least mobile of other figures to water, the most mobile to fire, and the intermediate to air. And again we assign the smallest figure to fire, the largest to water, the intermediate to air; the sharpest to fire, the next sharpest to air, and the least sharp to water. So to sum up, the figure which has the fewest faces must in the nature of things be the most mobile, as well as the sharpest and most penetrating'[34].

Even these snippets make the structure of Plato's reasoning clear. It is basically about the number of faces and the degree of slipperiness, from which he infers the different behaviours of different types of matter. It is a very practical and down-to-earth theory, with no whiff of symbolism. How architectural theory managed to go from here to the idea that 'Platonic' solids were spheres, circles, cones and cylinders, and the like – as is still repeated in many architectural primers on architectural form, too many to enumerate here – and as such constituted the visual and symbolic elementary forms of architecture is a mystery which some more diligent architectural historian may eventually unravel. It would be, at the very least, an instructive study in the persistence of error, and would certainly lead to the conclusion that at least some well-known names who talk about Plato had not troubled with the actual texts.[35]

The source of the original confusion might perhaps lie in the text where Plato actually does talk about geometry from an aesthetic point of view. It occurs in the *Philebus*, and even there his formulation is remarkably prosaic: 'Straight lines and circles and the plane or solid figures that are formed from them by turning lathes, rulers and measurers of angles . . . these I affirm to be not only relatively beautiful, like other things, but absolutely and eternally beautiful, quite unlike the pleasures of scratching.'[36] Turning lathes? Measurers of angles? Scratching? It is all a great way from mystical geometry and 'transcendental essences' that Pérez-Gómez, and others, have found in these forms.

In fact, a much more interesting and properly 'Platonic' suggestion is being advanced: that there is a connection between our knowledge of a form and knowledge of what we might call the genetic construction of that form. It brings to mind a much more modern formulation on the nature of perception, due to Jean Piaget: 'To know is to construct or reconstruct the object of knowledge in such a way as to capture the mechanism of that construction'[37]– a syntactic rather than a semantic theory of knowability.

I call this formulation Platonic because the essence of Plato's argument concerns how it is that the world is intelligible to us, that is to say, how the perception of the physical world is created by an *a priori* 'conceptual' world; once again, it seems that Pérez-Gómez has got it completely wrong – the whole force of Plato's argument about the mind and the world is to show the primacy of the conceptual over the perceptual. The Platonic theory of 'forms' is, in fact, about this intelligibility, and to think of them as 'transcendental essences' is simply an error.[38]

Plato's theory of form arises in the first place, from a very simple problem in everyday life. In language we use both *particular* and *universal* terms. A particular term is 'bird', as in 'that bird flies'. We may point to the bird and indicate it as we say its name. A universal term is 'birds' as in 'all birds fly'. 'Birds', in this case, cannot be indicated, since what is referred to is the class of all birds, whose locations are for the most part unknown to us. 'Birds' is therefore an abstract concept. One would suspect it of being a pure mental construct, were it not for the fact that it seems to depend on some abstract notion of a bird, which somehow sums up 'birdness', yet it is present in all real birds, regardless of their individual variation. 'Birds' seems to imply some kind of 'bird genotype'. These genotypes are, in effect, Platonic *forms*. We cannot argue that they are purely mental constructs, since they seem to depend on and interact with physical reality. Yet we cannot doubt that, in some senses at least, the bird 'type' is real enough, even if it exists as an abstraction.

If this is the case for birds, then we can easily imagine how much more difficult it is when we talk of circles rather than birds. What is a circle? Hardly a real one, since all of those must be imperfect. Is it knowledge of how to generate a circle? This won't do, because it says nothing about the final form which encapsulates at least part of our knowledge of circles. Is it purely a mental construct? Hardly, since circles are real enough. We have to accept that 'circleness' is somehow present in all these, but none has precedence in defining what a circle really is. We have to conclude, therefore, that the *theoretical possibility* of a circle is somehow at least as 'real' as any imperfect circles we might make, or any mental conceptions of them we might have. These abstract possibilities are, again, Platonic 'forms'. They are not mystical essences, but deductions from the dominant epistemological problem of our existence as experiencing subjects. Without such abstractions, the world would cease to be intelligible to us. Forms express the dominance of the perceptual by the conceptual and try to explain the intelligibility of things.

Plato's theory of knowledge is then really about three things: the idea of genotypical *forms*; the idea of the *genetic construction* of these forms; and *intelligibility*. It is clear that each of these ideas corresponds to one of the central interests of architectural theory. It is also clear that the Platonic concepts which have influenced architectural theory are not these, but naturalised mystifications of them; the mystification from the *Timaeus* and the symbolic mystification by the misreading of the idea of forms.

Is there a logic of architectural form?

What then follows for the understanding of the relation between theory and meaning in architecture today? Let us review the argument so far. First, I argued that the notion of meaning, in architecture as elsewhere, depended on pre-given theoretical constructs, and that structuralism was about the understanding of these constructs and how they permitted us to impose meaning on and derive meaning from artefacts. Then, I argued that architectural theory had always been structuralist in this sense, since it was centrally concerned with the relation between order and intelligibility. Unfortunately, I argued, the dominance of the mathematical paradigm of order had the effect of 'naturalising' the study of meaning, the latter being dominated by the linguistico-historical paradigm. In isolation from each other, neither branch of theory had progressed.

Next, I looked at two recent attempts to take advantage of the weakened state of 'structuralist' architectural theory, one to demolish it altogether (Scruton), the

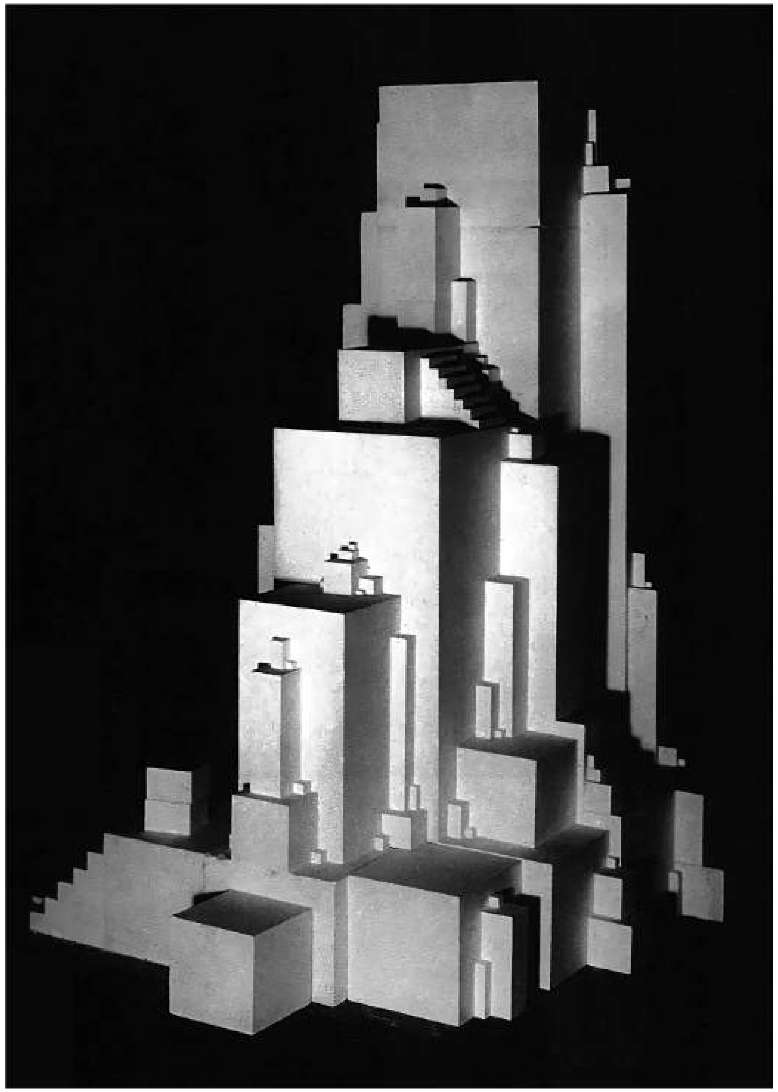


Figure 9.5 Malevich: Architecton Zeta 3D model.

other to force it to retreat into mysticism (Pérez-Gómez), both in the service of different philosophies. I have argued that both attempts have failed, the former because the analysis leads straight back to the structuralist problem of form, and the latter, because it leads us back to Plato, where we saw that both 'Platonic' ideas that had dominated architectural discourse – the idea of geometrical symbols, on the one hand, and the Timaen geometrical theory of matter, on the other – were illegitimate borrowings in the task of explaining order and intelligibility in architecture. On the contrary, it has been argued that each has constituted a different kind of mystification of the problem, the first a symbolic mystification, the second a naturalistic one.

But once these mystifications were cleared away, the central issues of architectural theory remain profoundly Platonic – in the real sense of the word; that is, how to describe order in form, both through its intrinsic nature and through its genetic construction, in such a way as to show how form carries intelligibility and meaning. Is there an alternative approach, then, one which dispenses with mystifications – both symbolic and naturalistic – but retains the notions of intelligibility, genotypical forms and genetic construction?

I suggest there is. The gist of my argument – which can only be roughly sketched here – consists of two propositions. The first is that the formal terms which we use to describe architectural form – proportion, symmetry, composition, rhythm and so on – do not rest on some unarticulated mathematical foundation, but are essentially *relational* terms. They refer simply to the variety of ways in which parts can be related to other parts, and to wholes. As such, their essential nature is logical rather than geometrical or numerical, and what we read as the intelligibility of form is, in the first instance, the *relational logic* or syntax of that form. The *significance* of form (that is, how architecture refers to itself) is thus, in the main, a matter of its relational syntax.

The second proposition is that *social signification* of form (that is, how architecture carries meaning) arises, in the first place, from the different types of elementary logical relations that are possible in the construction of architectural form. These elementary differences lead to different modes of constructing *complexity* in form, and these different modes will always carry with them the significatory potential of the types of elementary relation that dominates their syntax. The signifying potential of form will then depend on three factors: the *type* of relational syntax used; the *degree* to which syntax is used; and the degree to which *mass* is invested in particular syntactic relations – that is, a large mass invested with certain syntactic relations will read differently from a small mass invested with similar relations.

What these elementary relational ideas might actually entail can be explored by considering the common concepts which we use to describe form. Take symmetry, for example, in its simplest manifestation of bi-lateral symmetry, as in Laugier's primitive hut. Symmetry might itself be thought to be one of the 'elementary' concepts we are seeking. But it is not. It is a complex concept containing two quite distinct ideas. The first is the idea of *congruence*, meaning that the two components of a symmetrical object can, in principle, be overlaid on each other and found to be the same. It is important that the property of congruence is quite independent of space. Objects are or are not congruent regardless of where they are located in space in relation to each other. Congruence is thus a *logical* notion, rather than a spatial one.[39] The second idea covers what is omitted by the notion of congruence: that of a certain kind of *spatial relation*; namely that the relation of one component of the symmetrical object to the other – that is, of one sub-object to the other – is the same as the relation of that sub-object to the first. This seemingly tautological property is precisely what is meant by symmetry in mathematical relations: that is, that *a* is to *b,* as *b* is to *a*. Once again, this is primarily a *logical* idea, even though it is one which is realised through spatial relations. The spatial relation is one instance of the more general abstract relation, *a*:*b*::*b*:*a*.

Thus, the idea of bilateral symmetry, which to us (though not to Vitruvius) appears as a mainly geometrical concept, is also a rather complex syntactic idea, defining that to mean the logic of relations: parts of a whole that are congruent to each other are also in a spatial relation to each other in such a way that the relation of each to the other is identical. One might argue that the idea of 'sameness' or congruence is applied twice; once for the parts of the whole, and once for their relation.

What is the advantage of this definition? First, a logical description of form is a description of something like its 'genetic construction': how it is put together in such a way as to be intelligible to us. It refers at once, therefore, to the properties of the symmetric object itself – its own formal laws of construction – and to the mental operations by which we 'retrieve a description' of it.[40]

Second, once the logical components of the definition are separated, then each can clearly be used independently without any loss of intelligibility. For example, noncongruent parts can be placed in a non *a*:*b*::*b*:*a* relation, and remain readable as exactly that. In short, the logical ideas can be permutated and combined to form *logically* more complex and differentiated forms.

Third, we can see that using different combinations of logical relations would create different *types* of order, and at the same time, the degree to which the logical relation were applied to aggregates of parts would vary the *degree* of order which appeared to be present in that object. Thus, one could characterise the kind of 'organic' conglomerate form of, say, hill villages as possessing a low degree of this kind of order, while at the opposite pole we would find highly geometricised forms, such as Renaissance Ideal Towns, which possessed a strong degree of this kind of order.

Fourth, we can see the relations among the standard terms in use in architecture to describe form, and show that these relations are quite systematic. For example, the notion of proportion invokes the notion of metric congruence of parts without invoking the notion of particular kinds of spatial relations. Conversely, the notion of composition invokes the notion of certain kinds of spatial relation, without necessarily invoking the idea of congruent elements – although in practice, of course, either term might be used in such a way as to imply the other. Similarly, the notion of rhythm implies a certain periodicity of congruence, again without necessarily invoking particular types of spatial relation.

Finally, we may easily see, through this form of logical analysis, how it comes to be that this logical order is either a numerical or a geometrical one. Numerical, or modular, theories are always based on what Vitruvius called *analogia* or 'correspondence among measures',[41] and so appears easily as an arithmetical idea. But, as has been confirmed by many theorists, including both Vitruvius and Scruton, a violation of arithmetical accuracy is not a violation of architecture. This is simply because the relational, or syntactic, idea implied by modular theories can easily survive, and even gain from, metric adjustments. As Vitruvius says: 'The first thing to settle is the standard of symmetry (that is, the module), from which we need not hesitate to vary.'[42]

The nature of geometry in architectural form can, therefore, also be made clear. Forms are geometrical to the extent that certain kinds of congruence and certain kinds of spatial relation are used in order to clarify the structure of whole patterns at the expense of the identity of its parts. What we might call 'geometricity' in architectural form is not therefore the *basis* of order, but is, itself, a *special kind of order*.

But this would only be the case where the concept of geometry implied, first and foremost, a geometric relation among well-defined elements in a composite whole. The idea of a pure geometrical form as defining the overall shape of a building is quite different, and in some ways the opposite, since it implies that geometricity arises not from certain relations among constitutive elements, but, in all likelihood, from the lack of such relations, since that would limit the legibility of the pure geometric form of the whole. In this sense, the use of pure geometric forms can be seen as a means for eliminating the forms of social signification associated with the modalities of formal syntaxes.

If this is the case, then it is possible to discern a profound relation between the two dominant formal notions of Modernism in architecture: that form itself is *geometrical in a pure sense*; and that form arises from *function*. Both are quite clearly ways of *naturalising* and, in effect, desocialising form. The notion of function naturalises the *genesis* of form, that of geometry the nature of *form itself*. In this sense, the apparently contradictory combination of concern for architecture as sculpture and of architecture as utility are unified as a determination to break with the inheritance of the past by breaking the link with the constitutive social syntax of architectural form. Hence the fundamental – and I believe unviable – Modernist intellectual project of trying to establish determinative relations between geometry and function as a means to rational architecture.

Leaving aside the viability, or otherwise, of the intellectual project, the question for design remains: how far was this attempt to work *outside* the social logic of form effective? Is form thereby truly naturalised and desocialised? Or does it, through its often unintended reflections of the syntax of form, continue to transmit the social message embedded in its elementary logic? It would seem that the ordinary experience of form now confirms that the social logic of form is not avoided by an *intention* to naturalise it, but persists through the nature of form itself. Buildings are judged to be 'monolithic' or 'bureaucratic' or 'capitalistic' regardless of the sculptural intention of the designer. And if we cannot side-step the social nature of form, then it would seem that we have no alternative but to understand it.

Acknowledgements

The ideas in this article owe a great deal to my discussions with students over the past few years, including Patrizia Zanella and David Chou, Helen Tsoskounoglou, and Richard Burdett; but especially to my discussions with Andreas Kourkoulas.

Notes

1. R. Venturi, *Complexity and contradiction in architecture*, with an introduction by Vincent Scully (New York: Museum of Modern Art, 1966); A. Rossi, *The architecture of the city* (Cambridge, MA: MIT Press, 1985) (Italian original as *L'architettura della città*, 1966).
2. R. Scruton, *The aesthetics of architecture* (London: Methuen, 1977); A. Pérez-Gómez, *Architecture and the crisis of modern science* (Cambridge, MA: MIT Press, 1983).
3. The editors have numbered and labelled the figures, which were not labelled or mentioned in the original text.
4. A. Colquhoun, 'Form and figure', in *Essays in architectural criticism* (Cambridge, MA: MIT Press, 1981), pp. 190–202.
5. A useful review of the linguistic analogy can be found in P. Collins, 'The linguistic analogy', in *Changing ideals in modern architecture 1750–1950* (London and Montreal: McGill University Press, 1965), pp. 173–184. It is instructive that Collins, writing in the mid-1960s, can view the language analogy as being out of fashion, especially so since from that time it has totally dominated the discussion of meaning in architecture through the influence of semiology. See also, M. Krampen, *Meaning in the urban environment* (London: Pion, 1979).
6. R. Scruton, 'Has architecture an essence?', in *The aesthetics of architecture* (London: Methuen, 1977), pp. 37–70, is particularly relevant to his case against theories of architecture.
7. Scruton, *The aesthetics of architecture*, p. 269.
8. Scruton, *The aesthetics of architecture*, p. 234.
9. Scruton, *The aesthetics of architecture*, p. 227.
10. Scruton, *The aesthetics of architecture*, p. 69.
11. Scruton, *The aesthetics of architecture*, p. 235.
12. Scruton, *The aesthetics of architecture*, p. 235.
13. L. B. Alberti, *De re aedificatoria* [On the art of building]. Book IX, *Ornament to private buildings*, translated by James Leoni (Florence: Published by subscription, 1452) (1845 edition), p. 202.
14. Alberti, *De re aedificatoria*, p. 203.
15. Alberti, *De re aedificatoria*, p. 194.
16. A. Pérez-Gómez, *Architecture and the crisis of modern science* (Cambridge, MA: MIT Press, 1983). The title is adapted from E Husserl: *The crisis of European sciences and transcendental phenomenology*, and his *Phenomenology and the crisis of philosophy*. No specific references are, however, made to these texts by Pérez-Gómez, and the dates given are those of the English translations rather than of the original German texts. (Editors' note: Hillier recommends the reader look up a review of *Architecture and the crisis of modern science* by Robin Evans in this same issue of *9H*: 'Lost treasure: *Review of Architecture and the crisis of modern science* by Alberto Pérez-Gómez, MIT Press, 1983', *9H*, 7, 3–4).
17. Pérez-Gómez, *Architecture and the crisis of modern science*, p. 7.
18. Pérez-Gómez, *Architecture and the crisis of modern science*, p. 8.
19. Pérez-Gómez, *Architecture and the crisis of modern science*, p. 19.
20. Pérez-Gómez, *Architecture and the crisis of modern science*, p. 10.
21. Pérez-Gómez, *Architecture and the crisis of modern science*, p. 10.
22. Pérez-Gómez, *Architecture and the crisis of modern science*, p. 19.
23. Pérez-Gómez, *Architecture and the crisis of modern science*, p. 22.
24. Pérez-Gómez, *Architecture and the crisis of modern science*, p. 10.
25. There are, in fact, other senses in which phenomenology can be used. For example, the phenomenology of Ernest Cassirer has more in common with structuralism than with the phenomenology of Husserl. This is because Cassirer is concerned with the analysis of projections of consciousness into the real world by symbols, languages, myths and other artefacts, rather than with the allegedly irreducible content of the subjective consciousness itself. One must always distinguish such 'phenomenology of the object' (which, properly speaking, all formal theories of architecture are) from the more common and usually mystical 'phenomenology of the subject'.
26. E. Husserl: 'Phenomenology', *Encyclopaedia Britannica* (1927); revised translation by R. E. Palmer in *Husserl: Shorter works*, edited by P. McCormick and F. Elliston (Notre Dame, IN: Harvester Press, 1981), p. 22.
27. K. Popper, *The logic of scientific discovery* (London: Hutchinson, 1959; 3rd edition, 1968), p. 111.
28. Vitruvius, *The Ten Books on Architecture*, translated by Morris Hicky Morgan with illustrations and original designs prepared under the direction of Herbert Langford Warren (New York: Dover Publications, 1960).
29. Vitruvius, *Ten Books*, pp. 24–31.
30. Vitruvius, *Ten Books*, p. 73.
31. Vitruvius, *Ten Books*, pp. 252–253.
32. See note 22.
33. Vitruvius, *Ten Books*, pp. 168–69.
34. Plato, *The collected dialogues of Plato including the letters*, edited by E. Hamilton and H. Cairns, 2nd printing with corrections, Bollingen Series 71 (Princeton, NJ: Princeton University Press, 1963). This is the standard edition of the *Dialogues*, but in the main, quotations are taken from the translation by H. P. D. Lee (London: Penguin, 1977), p. 77.
35. One major exception is Reyner Banham in R. Banham, *Theory and design in the first machine age* (Cambridge, MA: MIT Press, 1960), where he refers throughout to the Phileban – rather than the Platonic – solids.
36. Plato, *The dialogues of Plato: Philebus*, translated and introduced by B. Jowett (London: Sphere Books, 1970), pp. 98–99. (Eds.: This is likely an allusion to a passage in the dialogues in which the pleasurable relief of scratching an itch is discussed.)
37. J. Piaget, *The mechanisms of perception*, translated by G. N. Seagrim (London: Routledge, and Kegan Paul, 1969), p. 356.
38. The best account of the theory of intelligibility is set out in Plato, *The Republic of Plato*, translated and introduced by F. M. Cornford (London: Oxford University Press, 1941), pp. 216–221.
39. This might be termed a 'transpatial' notion since it quite specifically relates objects to each other independent of space. For a fuller discussion on the concept of the 'transpatial', see B. Hillier and J. Hanson, *The social logic of space* (Cambridge: Cambridge University Press, 1984).
40. Hillier and Hanson, *The social logic of space*, p. 37.
41. Vitruvius, *Ten Books*, p 72.
42. Vitruvius, *Ten Books*, p. 175.

10 The nature of the artificial (1985)

Purpose, law, function, explanation. Introduction to 'The nature of the artificial'

Sonit Bafna

Bill Hillier's writings are suffused with themes, concepts and language fashionable in the scholarly literature of his time, and this can make the content of his papers seem dated. Yet this impression is misleading, for what gives these texts their staying power is Hillier's uncanny ability to find consequential insights buried within the debates of the day. In the 'The nature of the artificial' the core insight comes in the middle of the section titled 'Laws of the artificial'.

> I will try to show, by means of examples, that even in the case of spatial form in architecture, Aristotle's starting point on the notion of *purpose* is as devoid of explanatory power as it is in the natural sciences, for more or less the same reasons. The analysis of purposes is, of course, a vital part of the pragmatics of architectural understanding. But it cannot be the basis of a scientific understanding of spatial form in architecture.

Hillier develops his ideas as a response to Herbert Simon and invokes Aristotle, but the radicality of his thinking here is better understood in the context of mid-twentieth-century debates on the nature of scientific explanation and particularly on the established distinction between the explanation of actions (which necessitated invocation of purpose) and the explanation of natural phenomena (which called for a subsuming law). The history of this remarkable moment – yet to be written – would cover a fascinating variety of thought, ranging from provocative essays by critics like William Wimsatt and Monroe Beardsley, I. A. Richards, to lively debates such as the one between the historian William Dray and the philosopher Carl Hempel.[1] It would reach backwards to include Wittgenstein, whose students – Henrik von Wright and Elizabeth Anscombe – were key figures in establishing the framework of the debate, and forward to the work of Donald Davidson, who established the impossibility of psychological and psychosocial laws.[2] In all these writings, a central concern was the recognition that products of human action exhibited underlying regularities in their form that escaped intention but also resisted explanations as products of natural law.

It is something of a pity that Hillier did not pursue the themes of this paper further, for he would then have found himself confronting the same questions about the nature of law and human action. He would have seen that laws that he actually identifies as examples are an odd sort. The law exemplifying type 2 – 'inequality relations between function and spatial arrangement are stable across a sample' – turns out to be an empirical regularity, since there does not seem to be any reason to believe that it would *necessarily* hold for all samples of a building type. The related pairs of laws of type 3 seem to be only definitions. Even for the example of type 1, which *prima facie* does seem like a law ('the principle that a local constraint or rule applied to a random process would lead to a global form'), a claim might be made that the underlying systematicity is a mathematical theorem and lacks the axiomaticity of natural law. What, then, explains the regularities in built form if not natural law?

Fascinatingly, the practical solution to this issue lies in Hillier's work as well. Biologists and social scientists have argued since the mid-twentieth century that social phenomena or biological traits are best explained by *functions* rather than laws or purposes. To take a classic example, the claim that the function of the heart is to pump blood explains both its shape and its presence in the body of an animal. Similarly, to argue that the function of partitioning a building is to reorganise spatial proximities amongst residents and activities, or that the function of the integration core that spans a village layout is to ensure intelligibility, is to offer meaningful explanation of aspects of spatial organisations of built form. This is of course the central theme of 'What is building function?'

To put forth a functional explanation is not to deny that laws are involved at some level in determining human action, but rather to hold that laws act upon historically contingent factors, and that the focus of explanatory interest in the case of human constructive action is on these factors. It was characteristic of Hillier to have homed-in, pursuing his own line of thought, to the crucial point at which natural lawfulness confronts the deeply human element of the built environment.

The nature of the artificial

The contingent and the necessary in spatial form in architecture

Bill Hillier

Originally published: Hillier, B. 1985. 'The nature of the artificial'. *Geoforum*, 16, 163–178.

As an explanatory – as opposed to a descriptive – concept, *purpose* has long since been dispensed with in natural science. But recently it has taken up residence in the 'sciences of the artificial' since, it is argued, the form and nature of artefacts can only be understood in terms of the human purposes they serve. The aim of this paper is to show, by examples, that in the case of architecture and urbanism – the largest human artefacts apart from society itself – the argument is fallacious. Architectural and urban artefacts exist to organise and order space for human purposes. But their form and nature cannot be explained in terms of these purposes alone. Explanation requires an understanding of morphological laws of space, within which human purposes work themselves out. The laws are, it would seem, akin to natural laws, and would not justify any special epistemological stance for the 'sciences of the artificial'. *Purpose* should not therefore be given any more privileged status for artefacts than it has for nature. In both areas the role of science is to give an account of the underlying morphological constraints on possibility.

Aristotle's houses

It is often observed that Aristotle's philosophy of nature was based on the assumption that the existence of forms well-adapted to functions in nature was evidence of purposeful design. It is less often observed that his paradigm was architecture.[3] The form of a house, he argued, cannot be explained merely by the process of laying stone on stone, since this 'material cause' needs to be guided by a pre-existing idea of the form the house is to take. Its material form, we might say, needs to be guided by foreknowledge of the spatial form that is required. The argument is unexceptionable. Transferred to nature, it might reasonably be seen as an anticipation of modern genetics.

But Aristotle goes further. The problem, as he sees it, is not simply to explain form, but to account for the adaptation of form to purpose. His answer is immediate. Form and purpose are the same thing. Nature gives form in order to achieve purpose. As with architecture, as he sees it, form is incomprehensible without purpose. 'Final causes', therefore, are purposes. The source of all order in nature is purposeful design. On this architectural foundation the whole fallacious structure of Aristotelian 'science' was erected.

It had one merit. It was the philosophy in the rejection of which scientists would grow strong and wise. In a way, it set the standard. Any theory that could be shown to fall into the Aristotelian error of 'explaining' order in terms of some 'pre-existing order' – an essential requirement of a 'design' theory – could not be counted as a scientific theory. This is why such concepts as inertia in physics and random mutation in biology had such epistemological force, as well as scientific usefulness. They showed how it was possible, and necessary, to set theories of order on a foundation of non-order, thus finally breaking the link between the idea of order and the idea of design. With Aristotle, purpose had been an ultimate explanation. In science the relation of behaviour to form, which had given rise to the concept of purpose, became a problem of description. The concept of 'function', perhaps, embodies this descriptive neutrality.

Laws of the artificial

But the idea of purpose as an explanation of form is not dead. It is alive and well in the area where it began with Aristotle: the artificial. The views of Herbert Simon, inventor of the splendid concept of the 'sciences of the artificial', are pertinent here.[4] Artefacts, according to Simon, are unlike natural phenomena in that they are both contingent – we make them as we wish, rather than as nature makes them according to her laws – and purposeful – in that their form arises from the human purposes they are to serve. More precisely, Simon sees purposes (or 'goals', as he calls them) as links between two kinds of *environment*: the 'inner environment' of the artefact, that is its internal structure; and the 'outer environment', that is the range of circumstances in which these 'goals' must be realised. Purposes or goals are the means by which a selection is made from the contingent possibilities. Purposes are, therefore, the distinctive subject matter of the sciences of the artificial, in contradiction to both inner and outer environments, which belong in the natural sciences.

From this Simon derives a research strategy which can be summarised by his dictum adapted from Alexander Pope: '. . . the proper study of mankind is the science of design'.[5] By 'design' he implies a kind of means-end analysis for relating inner and outer environments as favourably as possible, given specified goals. Such a science is possible because:

> Just as the 'inner environment' of the whole system may be defined by describing its functions, and without detailed specification of its mechanisms, so the 'inner environments' of each of the sub-systems may be defined by describing the functions of that sub-system, and without detailed specification of *its* sub mechanisms.[6]

The 'sciences of the artificial' thus become coterminous with the 'science of design'. The two possibilities, Simon asserts, stand or fall together.[7]

Now this all sounds very convincing, until we realise that it has the rather bizarre implication that the sciences of the artificial do not exist, first and foremost, to study artefacts themselves, but to set up processes for designing them which reduce the artefact itself to an abstract representation by way of the notion of 'goal' or purpose. It is all the more bizarre if we reflect that, on a rather broader historical perspective, many sciences of the artificial are well-established and preoccupy themselves precisely with the inner structure – Simon's 'inner environment' – of the artefact. These sciences exist, by and large, where artificial systems are found with the curious property of being both man-made and not understood by man. Languages, cities and societies all come into this category. Moreover, in those areas where design plays a central role – such as urban design – it would seem that the simplification of the artefact itself to a functional description by the designers had been, historically, the means by which all understanding of the artefact was lost.

I have argued elsewhere that design depends on the availability of much more powerful representations of the artefact than would be given by a set of functional abstractions.[8] This is not my theme here. My argument in this paper is more fundamental, and concerns the nature of the alleged differences between artificial and natural systems. I will try to show, by means of examples, that even in the case of spatial form in architecture, Aristotle's starting point on the notion of *purpose* is as devoid of explanatory power as it is in the natural sciences, for more or less the same reasons. The analysis of purposes is, of course, a vital part of the pragmatics of architectural understanding. But it cannot be the basis of a scientific understanding of spatial form in architecture. The object of scientific understanding here, as elsewhere, must be the morphological laws and constraints within which architectural purposes work themselves out. Understanding the relation between purpose (or perhaps we should now use the term 'function') and spatial form depends on a prior understanding of something like the laws of space itself. Or, to put it another way, how function can turn itself into form is severely constrained by the nature of space itself, so much so that an understanding of these constraints and laws is necessary to any non-trivial analysis of the relation of function and form. This implies, of course, that in this particular realm of the artificial there exist morphological laws of the artefact itself which, while not being as powerful and universal as natural laws, are nonetheless sufficiently akin to them to justify an epistemological parallel between natural science and the sciences of the artificial.

Taking spatial form in architecture and urbanism as my starting point, I will present a series of simple examples to illustrate *three kinds of spatial law*:

1. Laws relating to the generation of spatial patterns by means of walls, apertures and so on, which we might call the *laws for the construction of space*.
2. Laws by which different types of social relations (for example, those which require space to provide segregation or integration) require different types of spatial pattern, which we might call *laws from society to space*.
3. Laws by which spatial structure has its effects back on society, which we might call *laws from space to society*.

The last of these raises the controversial question of 'architectural determinism'. One consequence of my argument is to show that, in a certain limited sense, such a thing does exist. Spatial design is neither empty nor trivial in its effects on society.

Each of the three types of law will be explained, in turn, by means of real cases. But because the interest here is primarily philosophical, technical explanation will be kept to the minimum necessary to make the cases clear. Readers interested in the full technical details should consult the appropriate texts.[9]

Type 1: Laws for the construction of spatial objects

The first example deals with the structure of certain small settlements. Figures 10.1a and 10.1b represent two small settlements in the north of England. Figures 10.2a and 10.2b represent two settlements in the south of France. All four are of the kind that would normally be classified as 'irregular' or 'random', since they all lack the kinds of defining features that govern standard classifications: village greens, linear streets and so on. Yet, oddly, their very irregularity seems to suggest that they might be – at some level – of the same *organic* type.

Careful description can indeed reveal certain common properties. Each settlement is composed of dwellings that define, by their location and orientation, a continuous structure of open space which has a very characteristic form. Firstly there is a continual variation in the width of the open space, so that the fatter areas of space are linked to other fatter areas through thinner spaces, creating an effect rather like irregular beads on a string. Secondly, the structure of open space forms at least one, and maybe several, intersecting rings, which means that there are always alternative routes from every point in the settlement to every other point. Thirdly, although the maps do not show this, empirical inspection will reveal that nearly every distinguishable piece of space – bead or string – is directly adjacent to at least one building entrance, so that the spatial pattern seems to be, in some sense, defined by building entrances.

In short, the four examples all have in common what we call, for want of a better term, a 'beady ring' structure, defined by the orientation of building

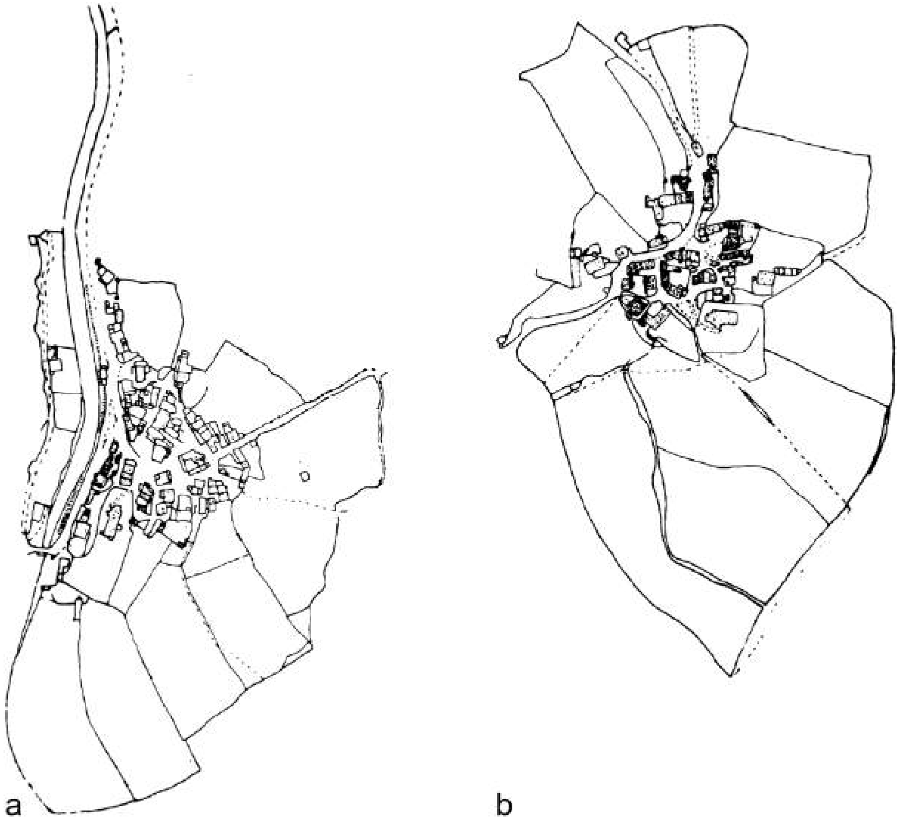


Figure 10.1 (a) Muker *(after OSC, 1888–1893)*; (b) Middlesmoor *(after OSC, 1888–1893)*.

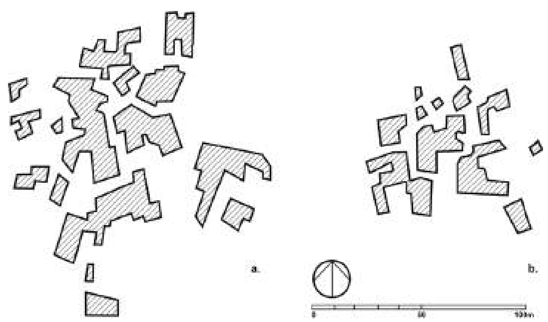


Figure 10.2 (a) Perrotet 1966; (b) Les Yves 1961.

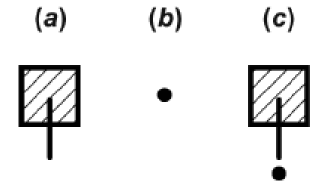


Figure 10.3 (a) Elementary unit with one entrance; (b) open space; (c) elementary unit with open space in front of entrance side.

entrances. What can explain this family likeness? None of the obvious candidates – climate, topography, defence, land tenure systems and so on – seem able to explain the specificity of the form as I have described it. So perhaps we should explore what might seem, on reflection, a more natural possibility than any of these: that the 'beady ring' form might in fact be the result of some simple process of *organic* growth.

A simple model can suggest such an aggregative process. Let us define an 'elementary unit' for this process as being made up of a simple square dwelling with one entrance and an equal-sized area of open space in front of the entrance side (Figure 10.3). Now, let us define a process of random aggregation of such units, subject to two rules: each new unit must be joined by its open space with a full facewise join onto an open space already in the aggregation; and while full facewise joins of dwellings are allowed where they occur randomly, the joining of dwellings by their vertices is not allowed (since it is unrealistic).

Figure 10.4 shows a fairly typical result of such a process of aggregation, carried out by computer and frozen at the average size of an actual 'beady ring' settlement. Figure 10.5 is a much larger surface generated in the same way. It is clear that all three defining characteristics of the beady ring structure – the formation of variable beads and strings of space, the formation of rings and the relation to building entrances – are all produced in a natural way by this process. Of course, one might object that the result is influenced by carrying out the process in an over-regular way. But this is not so. Figure 10.6 is a 'de-geometrised' computer-generated village layout (by Paul Coates of UCL) that simulates some rather more linear beady ring forms that can be found in the north of England. The general characteristics of the beady ring process are robust enough to appear in many variations.

What about the variation between the settlements in Figures 10.1 and 10.2? The main difference seems to lie in the size and number of clumps of buildings – the settlements in Figure 10.1 having several such clumps, and therefore several rings, while those in Figure 10.2 have a more dominant major clump and a dominant ring of open space. Again, a simple change in the formal model can generate this difference. In the process I described, the question of whether or not two 'houses' joined onto each other was settled randomly. If the process generated it, then

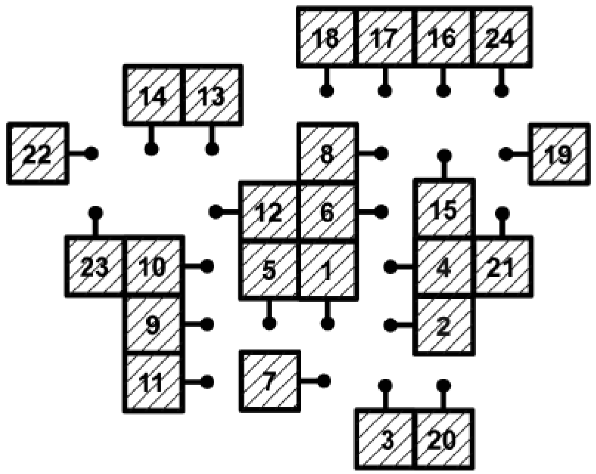


Figure 10.4 Computer-generated 'beady ring' surface. The numbers refer to the sequence of aggregation.

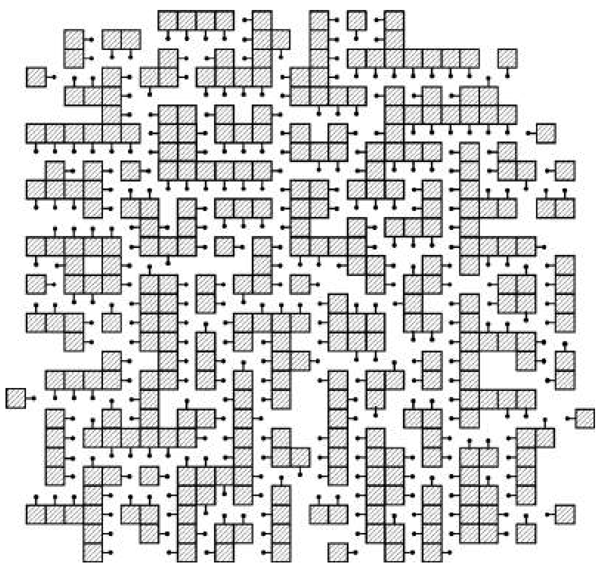


Figure 10.5 Large computer-generated 'beady ring' surface.

it happened. Suppose, then, we lessen the randomness of this type of join by specifying a higher probability of contiguous houses than would occur randomly. Obviously, the effect would be to increase the size and decrease the number of clumps. This is, in fact, the main difference between the settlements in Figures 10.1 and 10.2. If we continue to increase the probability of contiguous houses, we then eventually arrive at a pattern where much larger and more irregular clumps of

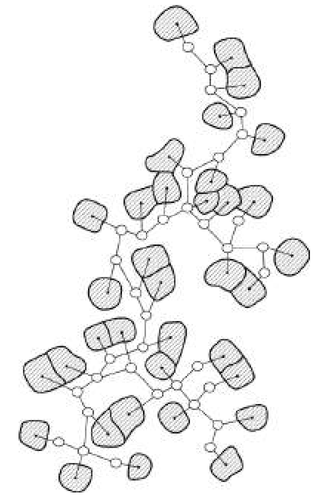


Figure 10.6 'De-geometrised' computer-generated village layout.

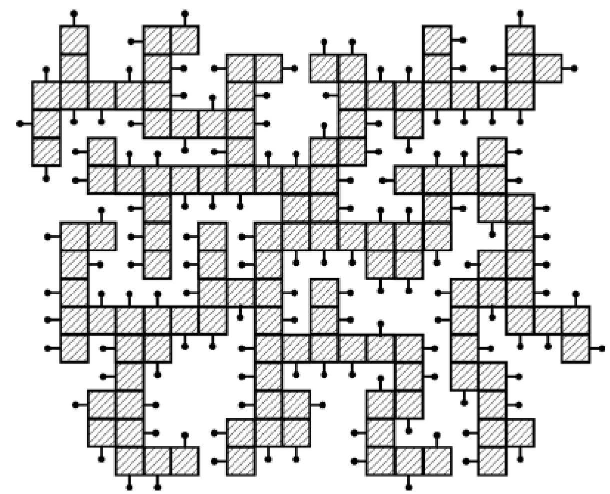


Figure 10.7 Computer-generated 'obverse' beady ring structure.

houses form deep wandering courtyards (Figure 10.7), as might occur, for example, in a settlement where building had taken place under severe spatial compression, say within a restrictive town wall. Figure 10.8 may well illustrate the end product of such a process.

Other variations follow from changing the *scope* of the rules. In Figure 10.9, for example, the basic beady ring process has clearly become more regularised by arranging the dwellings in such a way as to ensure that lines of sight (within the open space structure of the settlement) are more extended than they would

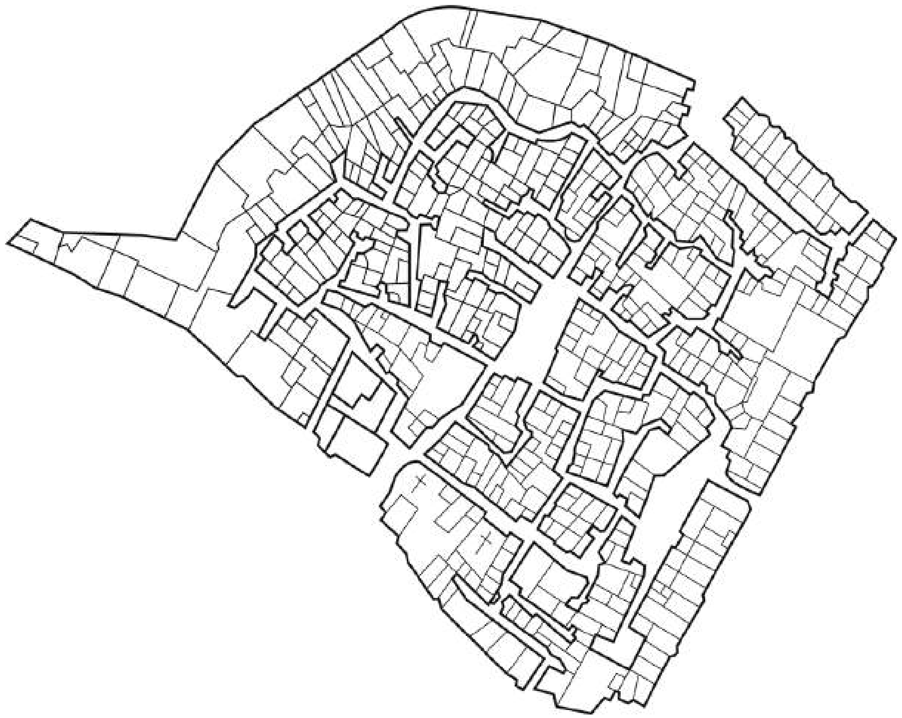


Figure 10.8 The walled town of Cisternino in southern Italy.

Figure 10.9 The town of Gassin in southern France.

be in a randomly generated aggregation. Such an effect could be easily produced by a rule that required each new unit to take into account not only how a house relates to its immediate neighbours, but also relations to other buildings at some distance away. It was, in fact, once suggested to me by a local builder in the area of this settlement that new houses could be added wherever one liked, provided that the new building did not block the view from any front door to any other existing door. Such a spatial rule would have exactly the effect of increasing the linear (or, perhaps, 'axial', to anticipate a later term) organisation of space.[10]

Can we say, then, that the process 'explains' the beady ring form? The question is difficult to answer on the basis of the completed forms, since clearly we need to consider the historical data on the step-by-step growth of each settlement. Such data are, unfortunately, hardly ever obtainable. We have, however, tried to test the proposition in a less direct way by mapping 77 apparently random clusters of buildings in the immediate vicinity of the French beady ring settlements. Figures 10.10a and 10.10b show two examples of two small clusters. Figure 10.11 is a larger one. Inspection of the sample shows first, that all clusters of sufficient size take a clear beady ring form, subject to topographical constraints, and second, that smaller clumps are to an overwhelming extent either compatible with a beady ring process of growth, or highly suggestive of it. Figure 10.10a demonstrates a 'compatible' case and Figure 10.10b a 'suggestive' case.

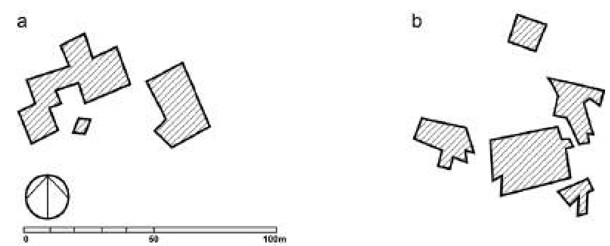


Figure 10.10 (a) Esquerade; (b) Les Bellots.

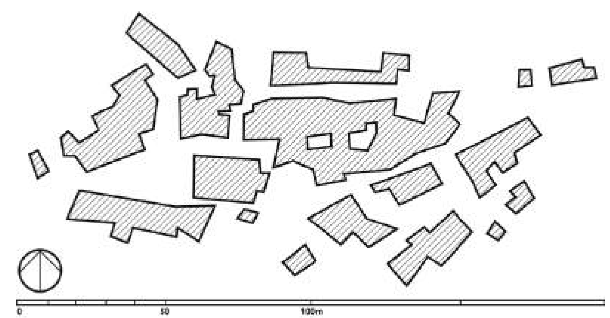


Figure 10.11 Les Petits Clements.

Suppose, then, that the beady ring process was involved in the development of these settlements. What would that tell us philosophically? Surely it would tell us that *two* kinds of knowledge, not one, were needed to 'explain' the settlement form. First, we would need the usual knowledge of real historical events and economic and social processes in which they were embedded. But this, on its own, would never truly 'explain' the special morphological invariants we have described. For this we need a second kind of knowledge; knowledge that included the principle that the local rule of 'open space contiguity' applied to an otherwise random aggregation process would lead *by itself* to the global form. This knowledge of the relation of implication between a local rule and a global form would be, at the very least, indispensable to a proper *explanation* of the form. In other words, we could not explain the form without knowledge of certain, quite autonomous 'laws of space' which, in this case, are to do with how some well-formed global properties could arise from the consistent application of a purely localised rule of aggregation.

This argument is taken forward in more extended texts to show both how a form of analysis of settlement space can be developed in order to identify 'genotypical' patterns in settlements (which appear to vary quite systematically with the type of society) and also to show how questions of social 'meaning' might be built into settlement forms by purely 'syntactic' means – that is, by means that have to do *only* with the organisation of space itself.[11] These considerations are too complex to be dealt with in a text of this length, but in any case, the idea that rules generating forms might have social meaning belongs to the second type of law.

Type 2: Laws from society to space

To illustrate laws of type 2 we can return to Aristotle's paradigm case: the spatial structure of the house. For this we need a little methodology: a representation and a measure. The representation employed is what we call a 'justified graph', while the measure is one of the degree to which a particular space *integrates* – or fails to *integrate* – a spatial complex.

Figure 10.12 is the ground floor plan of an English seventeenth-century house, one of a sample examined by the historian Wood-Jones.[12] The graph we are interested in is that of the relations of permeability – or access – from one space to another in the house. To make this graph we simply represent each space in the house by a small circle and each relation of direct access by a line, as shown in Figure 10.13a. To 'justify' this graph, we simply start from a point we are interested in – say, the space outside the house, as shown in Figure 10.13b – and align the graph up the page, so that spaces directly connected to the point of origin (the outside) are aligned immediately above this space on the first row, those two spaces away on the second row, and so on. This graph at once clarifies a property that we call *depth*. Depth exists to the extent that to go from one space to another

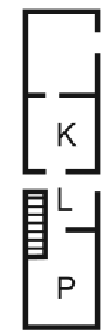


Figure 10.12 Ground floor plan of an English seventeenth-century house.

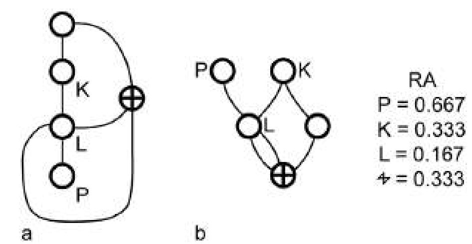


Figure 10.13 (a) Graph of permeability relations in the house shown in Figure 10.12; (b) 'justified' graph of the permeability relations from the space outside the house. The figures refer to the 'integration' values of the rooms.

it is necessary (that is, not merely possible) to pass through intervening spaces. We can say, then, that the space marked K, for example, is 'two deep' from the space marked P. Or that the space marked P is the 'deepest' from the outside. Or that the space marked L is the 'shallowest' on average from all other points.

Now consider the ground floor plans and respective justified graphs in Figures 10.14a–e. In each case P stands for parlour or best room; K for kitchen, or a space where food is prepared; and H for the space where most everyday living takes place. The justified graphs immediately show that all the statements I have made about the first case can be made about all these examples. P is always deepest from outside. K is never directly linked to P. H is shallower, on average, from all other spaces than either P or K.

Now, of course, it would be reasonable to expect that these 'invariants' are in some sense a product of the geometry of the house, or simple functional or

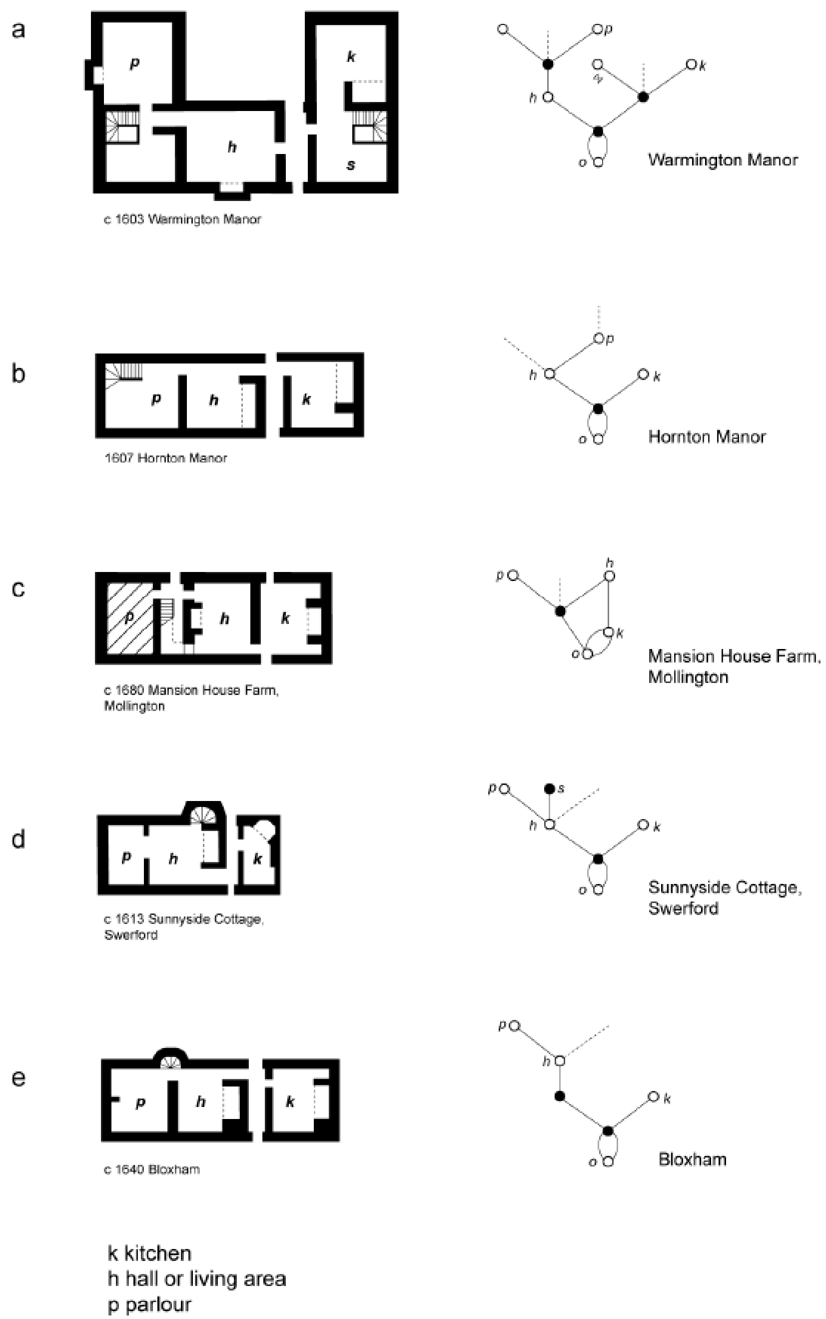


Figure 10.14 Five seventeenth-century regional house plans from the area around Banbury in England (from Wood-Jones 1963).

technological constraints of some kind.[13] So let us look at some more difficult cases. Figures 10.15a–d are a set of quite different English houses: a Victorian cottage, an up-market conversion of a Georgian house, a 1950s local authority house, and a lower ground floor flat conversion of unknown date. The positions of P, K and L have, it would seem, been juggled around quite considerably. For example, P is by the front door in two cases, on the top floor in another and the 'deepest' on the main floor in another. K is at the front in two cases and at the back in the others. Is there any underlying pattern in these cases?

To show that there is, let us first go back to the original house (Figure 10.12) and reverse our idea of the justified graph. Instead of looking at how 'deep' spaces are from the outside, let us begin from each space of interest – that is, P, K and L – and use the graph to show how deep all other spaces are from that particular space. Figure 10.16 shows justified graphs whose 'root' (starting point) is, respectively, P, L and K. These graphs show that, considered from the point of view of depth, each label has a different relation to the spatial complex as a whole. Other spaces are on average deeper from P than they are from K or L, and deeper from K than from L. Conversely, we can say that the complex, as a whole, is shallowest from L. Another way of saying this is that L *integrates* the complex more than other labelled spaces: integrates in the sense that it draws the complex closer to itself, whereas P, as it were, pushes it farthest away.

Now we have developed a way of measuring the degree to which any space *integrates* a complex in this way.[14] This allows comparisons to be made across complexes of different sizes, the 'mean integration' of different complexes to be calculated, comparisons to be made between individual spaces in different cases, and so on. Readers should refer to the above texts for further details. What is important here is that, on the basis of this measure of 'integration', we can make statements about the relation between room functions and spatial arrangements of the kind: L integrates more than K, which integrates more than P; or, to put it more briefly, $L>K>P$.

We can now return to the four apparently heterogeneous plans shown in Figure 10.15 together with the 'integration values' of the different spaces (plus the value from the outside and of the complex as a whole). Here we have to negotiate one little difficulty: a low value (that is, close to 0) indicates high integration and vice versa, since the value reflects the amount of depth the complex has from a particular space. It is then easy to see that in each case the following pattern exists: L integrates more than K, which integrates more than P; or $L>K>P$ for all cases. This is called an 'inequality genotype', meaning that certain inequality relations between the function and spatial arrangement are invariant across a sample.

What does this mean? It means, surely, that to the extent that such 'genotypes' can be shown to exist, then we have identified some kind of cultural principle for giving different social relations and activities a spatial form with respect to a whole complex of spatial relations. But if these *functions* are expressed through spatial invariants, then surely what we *mean* by a 'house' is a spatial complex in which these necessary differences can be related together in a way that preserves

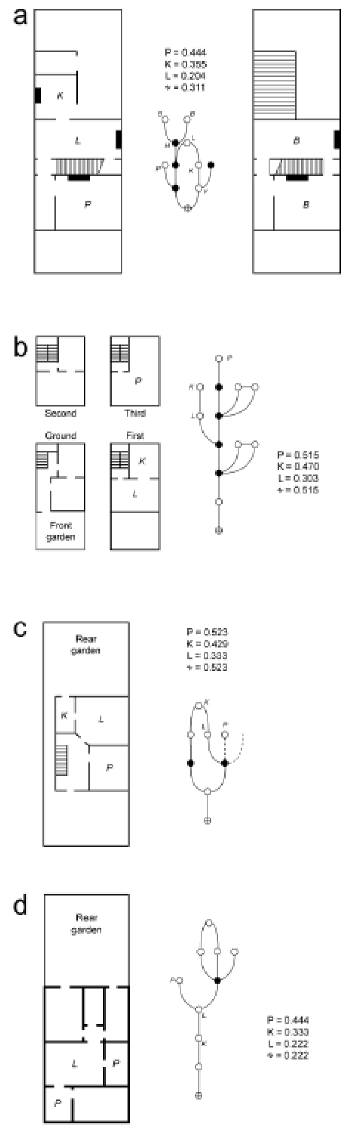


Figure 10.15 (a) Plans of a Victorian cottage; (b) plans of a converted house in Camden Town, London; (c) plan of a 1950s local authority house; (d) plan of a converted flat in Islington, London.

the 'genotypical' spatial relationship of each. Within the idea of an 'inequality genotype', the notions of functional patterns and spatial arrangement become, in a sense, the same thing.

Now we are *not*, of course, suggesting that the pattern we have identified is a cross-cultural pattern. On the contrary, we have presented a methodology for

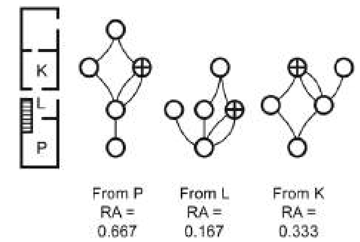


Figure 10.16 Plan of the house shown in Figure 10.12 with 'justified' graphs from P, L and K.

identifying one aspect of whatever cultural pattern exists (for other aspects of the methodology, again, readers should consult the referred texts). What is clear from our studies is that such patterns often exist and, while being markedly different from each other, nevertheless seem to follow certain underlying principles. Increasing the average depth from a particular space to all other spaces is a common means of achieving the relative segregation of important categories of spaces in buildings: front parlours, professor's rooms, chief's quarters and so on. On the other hand, varying the occurrence of 'rings' (or cycles) in the spatial arrangement is a normal means of varying the degree of control that particular spaces have with respect to the whole. Although cultures vary in their domestic space genotypes, the morphological means by which these genotypes are constituted are invariant.

Is Aristotle right, then, in this case? Does purpose give form? I would suggest in a descriptive sense not exactly, and in an explanatory sense not at all. In a descriptive sense it is clear that to follow the usual architectural convention of treating function *as though it were* a description of space is obviously inadequate. For example, to point to a space and say 'this is a kitchen' is clearly an inadequate architectural account of that space, since it omits any reference to the spatial relations of a kitchen in that culture. On the other hand, there could be a reasonable sense in which we might say that the 'purpose' of the spatial arrangement as a whole was to achieve a certain pattern of integration and segregation of functions.

But at an explanatory level such an account disregards the fundamental fact that the relation of form to function in such cases depends, firstly, on what space has to offer by way of combinatorial possibilities (that is, laws of type 1), greater or less average depth being one such possibility; and, secondly, on the fact that these combinatorial possibilities offer a family of *structural analogies* for different kinds of relationships among social categories and activities. For example, a 'semi-sacred' space like the 'best room' in a house can only be preserved as such if it has marginally greater *segregation* than the spaces of everyday activity. The social concept, as it were, already implies a 'genotypical' spatial requirement.

The relation of form to function thus passes through two types of morphological laws: those prescribing the lawful possibilities of combination in the first place; and those by which social relations give spatial meaning to such terms as ‘integration’ and ‘segregation’. Obviously, the task of an explanatory science must be to give an account of these types of morphological laws, not simply to *re-present* the problem by another functional description that takes these morphological preconditions for granted. All this can be made clearer, however, by looking at laws of type 3.

Type 3: Laws from space to society

To introduce laws of type 3 we will move from domestic space to the larger scale of urban space, since it is here that the *effect* of the pattern of space on its use by people can be most clearly discerned. Once again, we can make use of the notion of *integration* and its measure. Figure 10.17a is the plan of Barnsbury, an urban area in inner London, and Figure 10.17b is what we call its ‘axial map’, meaning the least set of straight lines that cover the open space of the area.[15] We might think of these lines as having to do with both what can be seen from a point and what is directly accessible from a point without changing direction. This axial representation – although it seems initially to be something of a reduction – has been shown by research to be important to both how people understand spatial patterns and how they use them.

Now it is clear that an axial map, like a set of rooms and doors, can also be thought of as a graph, with each line represented as a circle and each intersection as a line linking circles. We do not need to actually draw the graph to make an ‘integration’ analysis of the pattern formed by the axial lines. Clearly, each line will be either connected directly to another line, or it will be two lines or three lines deep from it and so on. Proceeding as before and treating every line in turn as though it were the ‘root’ of a justified graph, we can assign an ‘integration value’ to each line. As with domestic space organisation, these values will differ from each other, and these differences will be significant in understanding the nature of the spatial pattern and how the urban area works.

The first thing we can do is to identify what we call the ‘integration core’ of the area; that is, the 10 per cent most integrating spaces (or one could take 2 or 25 per cent, depending on the size of the area and what we want to show). Figure 10.18 shows the ‘integration core’ of the Barnsbury area in heavy black lines and numbered in order of integration. The most integrating line in the area is not, as one might expect, a long, central line but the rather short line that defines what is called the ‘village’, on which one finds a few shops, a pub and a garage.

Figure 10.18 also shows in dotted lines the spaces in the area that are more segregating than integrating; that is, the 50 per cent least integrating spaces. These quite clearly concentrate in the squares – a rather unusual pattern but one quite commonly found in nineteenth-century redevelopments. Note that although relatively segregated, these squares are still strongly connected to the surrounding urban fabric, and are for the most part ‘shallow’ from the spaces

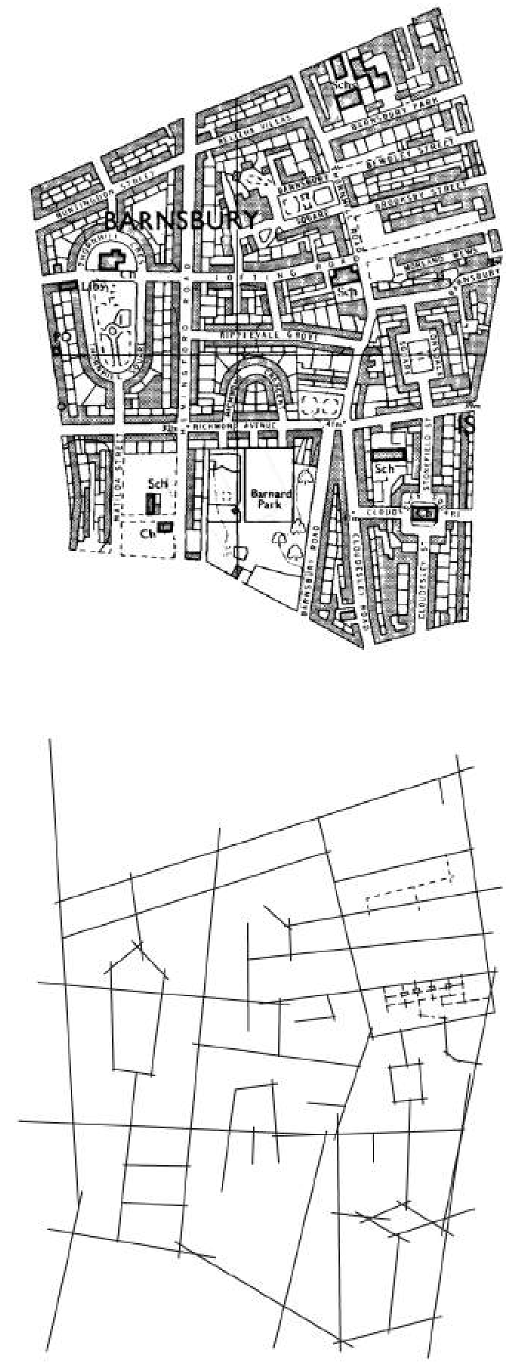


Figure 10.17 (a) Map of the area of Barnsbury in North London circa 1970; (b) Axial map of Barnsbury.

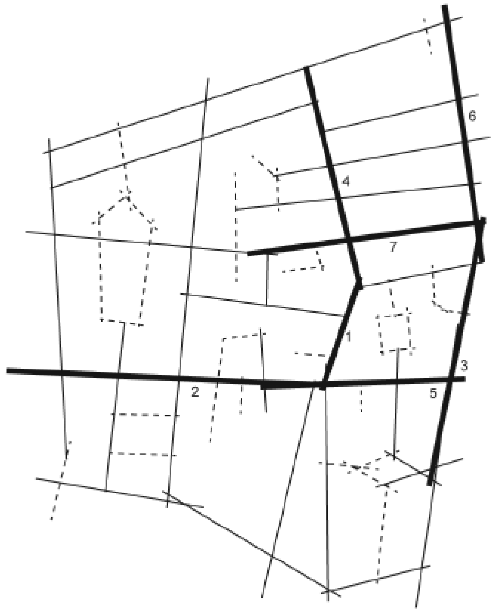


Figure 10.18 Integration map of Barnsbury.

on the periphery of the area that, in this case, is defined by the main roads. The arrangement of integrated and segregated lines does, however, show a rather common 'genotypical' pattern. The integration core links certain, fairly central lines to some peripheral areas in all directions. The core has, in effect, the shape of a deformed wheel, with a central hub and spokes radiating towards the rim and reaching it in several directions. A similar deformed wheel would have been found if we had subjected the settlement in Figure 10.9 to the same analysis, showing again that 'genotypical' patterns can be identified underneath considerable surface differences. Nonetheless, it must be emphasised that this particular pattern is far from being the only type that exists. On the contrary, the structure of the integrating core and the distribution of the more segregated areas are usually major clues to the type of culture that created the spatial pattern.

Why this might be so can be seen from Figure 10.19, which is a scattergram plotting integration against the numbers of people observed moving along different axial lines for a sample of spaces in the internal area of Barnsbury. I emphasise internal area because the peripheral spaces really form part of a larger system (also, however, with systematic properties at that level), and if they are taken into account as part of this local area, they tend to obscure the clarity of the relationship between spatial structure and patterns of movement. This relationship for the interior of the system is quite clear: there is a very powerful correlation between integration value and the tendency for people to use that line for pedestrian journeys.[16]

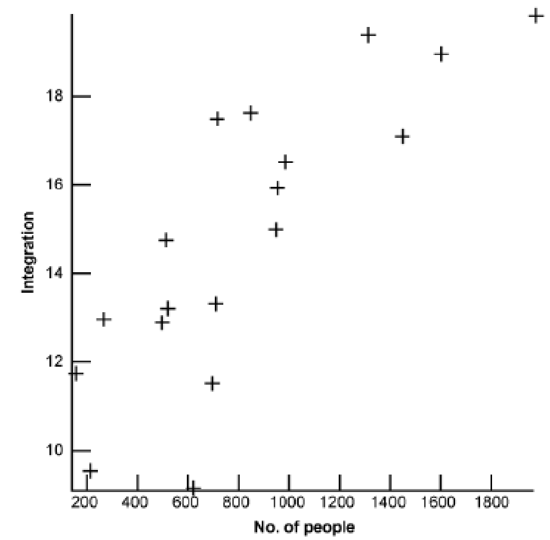


Figure 10.19 Scattergram plotting integration against the numbers of people observed in the Barnsbury area.

The curious thing is that if we then try another measure, one which measures the degree to which each axial line features on all shortest (that is, fewest axial steps) journeys from every point to every other point in the area, and which therefore should give the ideal distribution of moving people in the system, then we get a weaker correlation. Integration (rather than 'choice', which is the name we give to this second parameter) is the strongest predictor of movement, and this result can be duplicated for several urban areas that have been analysed and systematically observed in a similar way.

The reason is not hard to find, and it has to do with the shape of the integration core. The reason integration predicts movement so strongly is because many of the people observed are not moving only within the system but across or through it from origins or to destinations outside the system. However, the shape of the core makes it clear that the structure of integration in the system is arranged in such a way to facilitate movement from outside the area as far as possible. The core, by linking centre to edges in all directions, makes it easy for 'strangers' to enter and cross the system, while controlling it by making sure that the axially shortest journeys will concentrate on the most integrating lines. At the same time, the distribution of the core with respect to the interior of the system ensures that the more segregated destinations in the system are nevertheless in all cases 'shallow' from the integrating core.

At the same time, the existence of a strong correlation between 'choice' (which governs internal movement patterns) and integration (which governs global movement patterns into and across the system) ensures that a continuous *interface* is maintained between strangers entering the system and 'inhabitants' who are already within it. In other words, integration and other properties of the spatial patterns are used to produce that natural and unforced field of potential encounter between inhabitants and strangers – as well as between inhabitants themselves – that seems to be the hallmark of many traditional urban street patterns. It need, perhaps, hardly be added that it is exactly this global structuring of urban space, with its concomitant effects on the encounter field of people in the system, that is so often lost in pathological cases of modern urban estates and redevelopments.[17]

Is this effect from space to people an example of a 'purposive' artefact? Again, in a trivial sense we might say it is, since the 'purpose' of the arrangement of space seems to be to structure the pattern of movement. But once again this does not *explain* the nature of the pattern itself. As with beady rings and domestic space, the effect of space on movement depends on pre-existing morphological laws that relate space to movement, and it must be these laws which are the main objects of an explanatory theory.

In fact, the ways in which real environments work – either well or pathologically – are as often as not the *unintended* by-products of declared 'purposes' which designers have attempted to realise in a way that disregards these morphological constraints and laws. Most typically, designers have tried to design 'community' by using exactly those morphological means that most limit the possibilities of its development.[18] It may well be that history will eventually judge modern architecture in terms of this disjunction between purposes and realities – a disjunction that arises in the main from a concept of design in which a language of purposes and functions obscures the need for objective morphological knowledge.

Conclusion

So where does this leave Aristotle? In a way, my analysis confirms the distinction between a *material* process (or 'cause', as he would call it) and an *immaterial* one, since in all three types of cases that I have considered, something like abstract rules appear to intervene in a material process. But regularities in spatial form and socio-spatial behaviour do not arise *only* from those rules, but from their conjunction with the independent laws of material processes. Moreover, the rules themselves have a 'social logic' in that they are physical analogues to social ideas and relations. The relation of rules to forms is, therefore, neither direct nor simple. It passes through autonomous material laws, without a knowledge of which, any explanatory understanding of the situation is probably impossible.

Nor is it the case that these rules are, in themselves, akin to purposes. On the contrary, they are the *means* by which social processes can be expressed in spatial form. Without them, the relation of function to form must necessarily be arbitrary. In a sense it is neither function nor form which should be the object of an explanatory theory, but the means by which one turns into the other. It follows that the attempt to analyse spatial forms in terms of purposes can never be more than a superficial commentary on the obvious. So far as architecture and urban design are concerned, therefore, the language of purposes can offer little more to the 'sciences of the artificial' than it did to the natural sciences. 'Sciences of the artificial' must, after all, pursue the more orthodox path of uncovering the laws that govern the artefacts themselves – the laws of the artificial.

Notes

1. W. K. Wimsatt and M. C. Beardsley, 'The intentional fallacy', *The Sewanee Review* 54 (1946): 468–488; I. A. Richards, 'How does a poem know when it is finished?', in *Parts and Wholes*, edited by D. Lerner (New York: Free Press of Glencoe, 1963), pp. 163–174; W. Dray, *Laws and explanations in history* (Oxford: Oxford University Press, 1957); and C. G. Hempel, 'The function of general laws in history', *The Journal of Philosophy* 39 (1942): 35–48; C. G. Hempel, *Aspects of scientific explanation* (New York: Free Press, 1965).
2. D. Davidson, 'Psychology as philosophy', in *Philosophy of psychology*, edited by S. C. Brown (London: Macmillan, 1974), pp. 41–52.
3. Aristotle, *The basic works of Aristotle. Edited and with an introduction by Richard McKeon* (New York: Random House, 1941). The two most important references to the 'architectural paradigm' in Aristotle are probably the *Physics*, Book 2, Ch. 8, pp. 250–252 in the McKeon edition; and the *Parts of animals*, Book 1, Ch. 5, and Book 2, Ch. 1, pp. 657–659, also in the McKeon edition. But the idea is pervasive throughout his writings, as shown by the conceptual importance given to it in the two references cited.
4. H. Simon, *The sciences of the artificial* (Cambridge, MA, and London: MIT Press, 1969).
5. Simon, *The sciences of the artificial*, p. 83.
6. Simon, *The sciences of the artificial*, p. 73.
7. Simon, *The sciences of the artificial*, p. xi.
8. B. Hillier and A. Leaman, 'How is design possible?', *Journal of Architectural Research* 3 (1974): 4–11.
9. B. Hillier, J. Hanson, J. Peponis, J. Hudson and R. Burdett, 'Space syntax: A different urban perspective', *The Architects' Journal* 178(48) (30 November 1983): 47–63; and B. Hillier and J. Hanson, *The social logic of space* (Cambridge: CUP, 1984).
10. This approach to the understanding of morphology seems exactly analogous, philosophically, to René Thom's adaptation of C. H. Waddington's 'chreodic' processes, as set out in R. Thom, 'Structuralism and biology', in *Towards a theoretical biology*, edited by C. H. Waddington (Edinburgh: Edinburgh University Press, 1972), pp. 68–82. The analogy between approaches to describing natural and artificial morphologies illustrates perfectly the central thesis of this paper.
11. Hillier and Hanson, *The social logic of space*.
12. R. B. Wood-Jones, *Traditional domestic architecture of the Banbury region* (Manchester: Manchester University Press, 1963); P. Steadman, *Architectural morphology* (London: Pion, 1983), Ch. 12.
13. In fact, the difficulty of explaining these spatial features by technological or other factors is explored in J. Hanson and B. Hillier, 'Tradition and change in the English house', Working Paper of the Unit for Architectural Studies, University College London (1979); and B. Hillier and J. Hanson, 'Discovering housing genotypes', Working Paper of the Unit for Architectural Studies, University College London (1982); and B. Hillier and J. Hanson, 'Space after modernism', *9H* 3 (1982): 15–20.
14. See Hillier et al., 'Space syntax'; and B. Hillier and J. Hanson, *The social logic of space* (Cambridge: CUP, 1984), chapter 3.
15. Hillier et al., 'Space syntax'.
16. These results are based on research carried out in 1983 by the Unit for Architectural Studies at the Bartlett School of Architecture, which involved the systematic observation of patterns of pedestrian movement and occupancy in the Barnsbury area.
17. See Hillier et al., 'Space syntax'.
18. See, for example, the ideas and layouts presented in C. Alexander, S. Ishikawa, M. Silverstein, M. Jacobson, I. Fiksdahl-King and S. Angel, *A pattern language: Towns, buildings, construction* (New York: Oxford University Press, 1977).

11 Ideas are in things (1987)

The social logic of dwellings. Introduction to 'Ideas are in things'

Luiz Amorim

In the opening chapter of their first book *The social logic of space*, Bill Hillier and Julienne Hanson introduce what they believe to be the fundamental conundrum concerning the concept of space both in the context of the discipline of architecture, but also in the social sciences. They suggest that the very nature of the space-society phenomenon should be redefined based on a simple, but paradigmatic, argument: 'Society can only have lawful relations to space if society already possesses its own intrinsic spatial dimension; and likewise, space can only be lawfully related to society if it can carry those social dimensions in its very form'.[1]

To support this radical shift in the space–society paradigm, it was necessary to demonstrate how spatial arrangements of discrete elements, regardless of their size and purpose, allow the continuity or evolution of the social forms they house, as observed historically. It is argued that societies operate by means of a description retrieval mechanism that permits the recognition, description and subsequent application of organising principles directly from the spatio-temporal reality and activity. They evoke the idea of *inverted genotype*, whereby the informational structure does not reside within the genetic make-up of individual beings – the biological 'phenotypes', but within the human environment.

Indeed, the concepts of description retrieval and inverted genotype were key to unravelling how space operates as a relational system that embodies social purposes. The 1987 paper 'Ideas are in things', co-authored with Julienne Hanson and Hillaire Graham, is important because it introduces a rich analytical model for the inference of the genotypical and phenotypical properties of spatial systems, further developing ideas originally proposed in *The social logic of space*.[2]

The proposition is that the pattern of connectivity of spatial configurations defines asymmetries relative to space-to-space accessibility. The more central or integrated spaces are closer to all others. By contrast, the less central spaces are deeper in the system in that more intervening spaces must be crossed in order to reach them. The *mechanism* that assigns spatial dimensions to social functions is precisely this asymmetry of spatial systems.[3] Recurrent questions related to the social life of buildings, such as who does what, with whom, when and where, are thereby anchored in the configuration of space and so can be described mathematically. Put in simple terms, it is by means of the configuration of discrete spatial structures that social processes are expressed in space, and description is retrievable allowing for a self-reproduction of the system.

The consistent asymmetrical relationships between key functions, expressed in the rank order of spatial integration within a set of culturally significant buildings is named an *inequality genotype*, 'the most general means by which culture is built into spatial layout'.[4] The more differentiated the integration values are, the stronger the expression of genotypical patterns in the layout itself.

The investigation from which the inequality genotype concept was built was initially motivated by discussions with Professor Jean Cuisenier, Head Curator of the Museé national des Arts et Traditions populaires, over the possibility of analysing patterns of domestic life based exclusively on unidentified plans of a set of French farmhouses – *les maisons rustiques*, properly labelled, denoting domestic activities that take place in each constituent space. Domestic activity is understood to be a set of informal (namely a constitutive part of domestic daily life) and formal events (in circumstances in which codes of conduct are more precisely defined). Informal domestic ceremonies are carried out according to short behavioural instruction models, whereas formal events necessitate longer ones.[5]

Two inequality genotypes were identified from within the set of studied *maisons rustiques*. The most pervasive is centred at the space of everyday family use, where the reproduction of habits and customs of rural community takes place. The second has its configurational centre in a transitional space, whose effect is to segregate the other functional spaces and their users. On the basis of this and subsequent cross-cultural studies, inequality genotypes centred in functional and transitional spaces were recognised as universal ordering principles.

'Ideas are in things' presents a key to understanding, as Hanson stated, 'the ways in which people's dwellings embody and express cultural lifestyle preferences'.[6] It is one of the inaugural papers of the field of study dedicated to the social logic of dwellings, that culminates in Hanson's essential *Decoding homes and houses*, a book recommended to those interested in the relationship between house form and culture, but also a book that elucidates and develops further the ideas formulated in this ground-breaking paper.

Ideas are in things

An application of the space syntax method to discovering house genotypes

Bill Hillier, Julienne Hanson and Hillaire Graham

Originally published: Hillier, B., Hanson, J. and Graham, H. 1987. 'Ideas are in things: An application of the space syntax method to discovering house genotypes'. *Environment and Planning B: Planning and Design*, 14, 363–385.

Introduction to space syntax

Space syntax is a set of techniques for the representation, quantification and interpretation of spatial configuration in buildings and settlements.[7] Configuration is defined in general as, at least, the relation between two spaces taking into account a third, and, at most, as the relations among spaces in a complex taking into account all other spaces in the complex. Spatial configuration is thus a more complex idea than spatial relation, which need invoke no more than a pair of related spaces. The theory of 'space syntax' is that it is primarily – though not only – through spatial *configuration* that social relations and processes express themselves in space.

The primacy of configuration in the 'social logic' of space does not just *happen* to be the case. It originates in the logic of space itself. This can be simply demonstrated. Figure 11.1 is a divided cell in which space *a* is linked to space *b* through a gap. The gap creates a 'relation' – we might call it 'permeability' – between the two spaces. But it means little until we know the relation of each to at least one further space – that is, until we know the position of each with respect to a *configuration*. For example, Figure 11.2 shows two possible relations of spaces *a* and *b* to the outside, space *c*. In Figure 11.2a, both spaces are directly connected to *c*, but in Figure 11.2b only space *a* is so connected, so that it is necessary to pass through space *a* to get to space *b* from space *c*. This means that the relation between *a* and *b* is changed when *c* is considered. In one case, *a* controls the path from *c* to *b*; in the other, this is not the case.

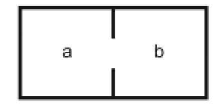


Figure 11.1 A divided cell.

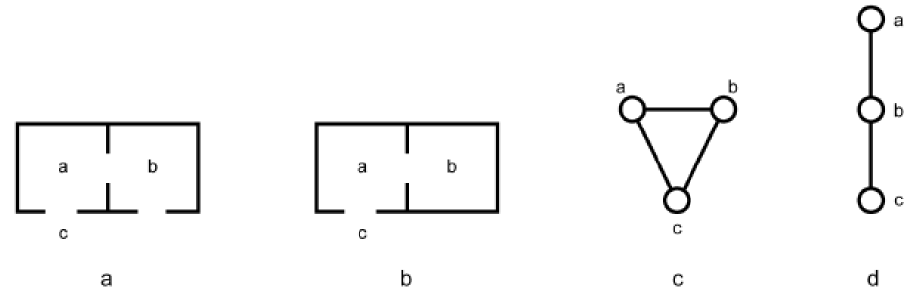


Figure 11.2 Two possible relations of spaces (a and b) to the outside (c) and the corresponding graphs.

This difference may be graphically clarified by a useful technique for representing spatial configuration: the *justified graph*. This is a graph in which a particular space is selected as the 'root', and the spaces in the graph are then aligned above it in levels according to how many spaces one must pass through to arrive at each space from the root. Thus Figures 11.2c and 11.2d are justified graphs of Figures 11.2a and 11.2b, respectively.

These two graphs also serve to illustrate the two configurational properties of spatial layouts which seem most important in articulating cultural ideas and social relations. The first of these is the property of *depth*. A space is at *depth* 1 from another if it is directly accessible to it, at *depth* 2 if it is necessary to pass through one intervening space in order to move from one to the other, at *depth* 3 if a minimum of two spaces must be passed through, and so on. In the justified graphs, therefore, depth from one space to another will show as height when the first space is used as the root.

The second property is that of *choice*, that is, the existence or otherwise of alternative routes from one space to another. Regardless of depth, all graphs which are *trees* – that is, those which have *k* spaces and *k-1* links – will have only one route from any space to any other. Alternative routes will therefore show themselves as *rings* in the graph, as in Figure 11.2c. Spaces can be distinguished from each other according to whether or not they lie on rings, how many rings they lie on, and which rings they lie on.

These two concepts will underlie all that is said in the following analysis. The first, *depth*, will, however, be used in a more developed and quantitative form which we call *integration*. The *integration value* of a space expresses the relative depth of that space from *all* others in the graph through the formula below, where $\bar{d}$ is the mean depth of spaces from the space and k is the total number of spaces in the graph. This gives a value varying between 0 for maximum integration, that is, no depth (as in Figure 11.2c) and 1 for maximum segregation, that is, maximum possible depth (as in Figure 11.2d).[8] The integration value of a space thus expresses numerically a key aspect of the shape of the justified graph from that space:

$$\textit{integration value} = \frac{2(\bar{d}-1)}{k-2}$$

In most spatial complexes, integration values will be different for different spaces, and justified graphs will show this difference visually. Figures 11.3a and 11.3b, for example, are justified graphs of the same complex drawn from two different points. Figure 11.3a is relatively deep, with an integration value of 0.43 or 1.31 using the transformation given by Hillier and Hanson,[9] whereas 11.3b is very shallow, with a value of 0.09, or 0.29 using the transformation.

Such differences are one of the keys to the way in which culture and social relations express themselves through space. For example, different *functions* or *activities* in a dwelling are usually assigned to spaces which integrate the complex to differing degrees. Function thus acquires a spatial expression which can be assigned a numerical value. If these numerical differences in functions are in a

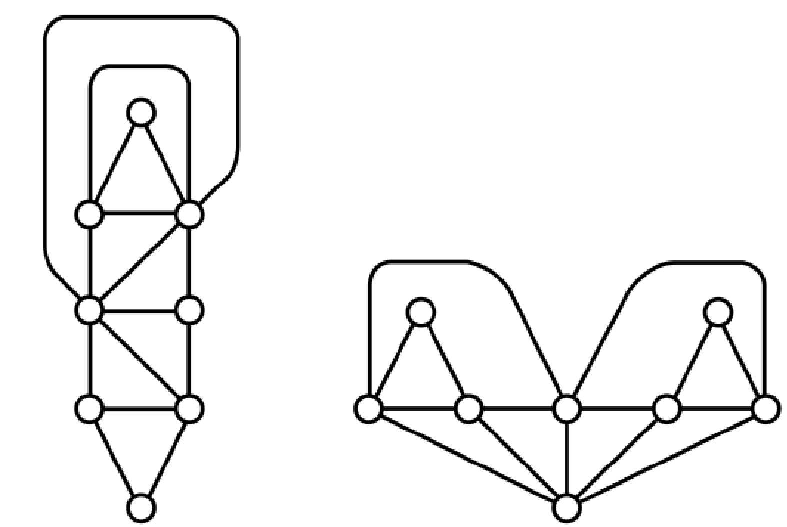


Figure 11.3 Justified graphs of the same complex drawn from two different points.

consistent order across a sample, then we can say that a cultural pattern exists, one which can be detected in *things*, rather than just in the way it is interpreted by minds.

This particular type of consistency in spatial patterning we call an *inequality genotype*. We believe it to be one of the most general means by which culture is built into spatial layout. How strong or weak these inequalities are in a complex, or in a sample, is therefore also of importance. To measure this, we have developed an entropy-based measure called *difference factor* to quantify the degree of difference between the integration values of any three (or more, with a modified formula) spaces or functions. This is essentially an adaption of Shannon's H-measure for transition probabilities,[10] in which we substitute the integration value of a space over the total integration for the three spaces for the transition probabilities in Shannon's equation, where H is the unrelativised difference factor for three spaces, a, b and c are the integration values of the spaces, and t is their sum:

$$H = -\sum\left[\frac{a}{t} ln\left(\frac{a}{t}\right)\right] + \left[\frac{b}{t} ln\left(\frac{b}{t}\right)\right] + \left[\frac{c}{t} ln\left(\frac{c}{t}\right)\right]$$

This H can then be 'relativised' between ln2 and ln3 to give a 'relative difference factor', H^*, between 0 (the maximum difference, or minimum entropy) and 2 (the minimum difference, or maximum entropy, that is, all values are equal):

$$H^* = \frac{H - \ln 2}{\ln 3 - \ln 2}$$

This relativisation is possible because the maximum H for k values is always ln k (in this case, therefore, ln3), and in the case of the integration measure, if one space has a value of 0, then it follows that the other two spaces must have a value of 1, in which case H is ln 2, and this is the minimum possible. To give the feel of this measure, the difference factor for, for example, 0.4, 0.5 and 0.6 is 0.97 (that is, close to 1 or very weak), whereas that for 0.3, 0.5 and 0.7 is 0.84, or considerably stronger, and that for 0.1, 0.5 and 0.9 is 0.39, or much stronger still.

These simple measures are, we believe, able to express culturally significant typological differences among plans because the two concepts on which they are based have in themselves a kind of intrinsic 'social logic'. *Depth* among a set of spaces always expresses how directly the functions of those spaces are integrated with, or separated from, each other and thus with how easy and natural it is to generate relations among them; whereas the presence or absence of *rings* expresses the degree to which these relationships are controlled, or marked by an absence of choice, forcing permeability from one space to another to pass through specific other spaces. Figure 11.4 shows typical, easy-to-remember patterns with these characteristics.

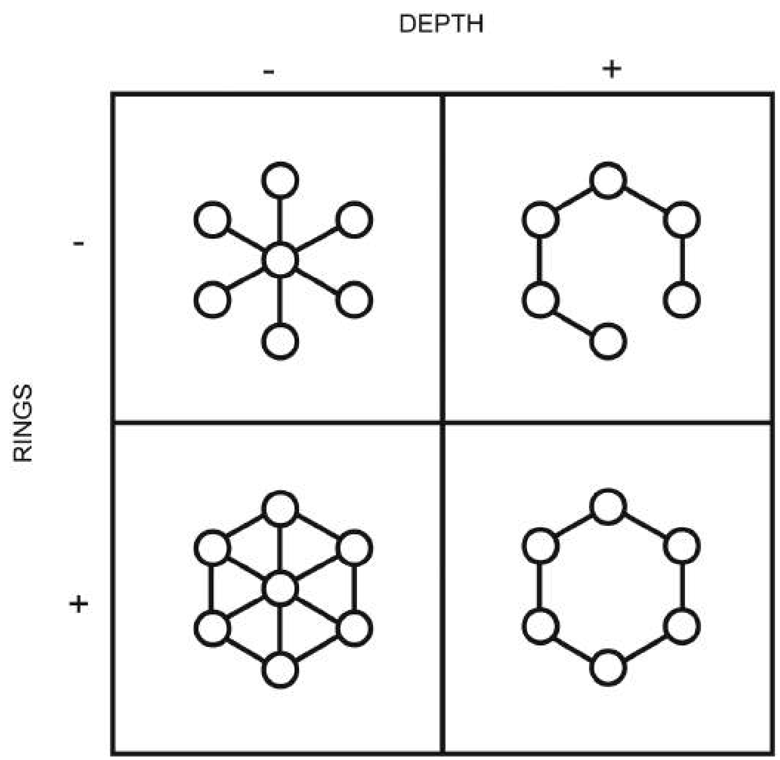


Figure 11.4 Typical patterns with differing characteristics of depth and rings.

It must be emphasised, however, that it is the *quantitative picture* of a spatial configuration that betrays its cultural bias, not a simple qualitative diagram. In the case of housing layouts, ideas are built into things, not so much through visual representation, as through the configurational principles by which a spatial pattern is constituted by its makers so as to be culturally *intelligible*. It is the hypothesis of this paper that these principles are capable of quantitative expression.

The sample and the problem

In June 1985, Professor Jean Cuisenier, Conservateur en Chef du Musée national des Arts et Traditions populaires, after discussions in London, sent samples of plans of rural dwellings from various regions of France to the Unit for Architectural Studies[11] for spatial analysis by means of the 'space syntax' method. It had been agreed that the analyses should initially be carried out 'blind', with no information apart from the plan labelled with its various functions – one might say, with no more information than an archaeologist might have. Social, economic and cultural information would be explored only after the initial spatial analysis. The question was: how far was it possible to analyse domestic space patterns with only *artefactual* or *archaeological* information? This paper is concerned only with this initial spatial analysis, and socio-cultural issues are raised at the end only in the form of speculations requiring further – and a different kind of – research. Within these restrictions, the aims of the analysis were:

1. to see how far syntactic representations and analyses could clarify the relation between patterns of space and their use;
2. to ascertain how far regional or other types might be suggested by such an analysis;
3. to explore the possibility that certain known traditional themes might be reproduced in at least some of the houses, and that these themes might be clarified by syntactic analysis.

These traditional themes were derived from Cuisenier's exposition of *La maison rustique* by Charles Estienne, published in 1564.[12] In his exposition, Cuisenier proposed that Estienne's account of the 'maison rustique' (as opposed to the château or the manorial domain) could be clarified by an underlying model with three elements (Figure 11.5): *orientation*, regulating the general orientation of the farm and its built elements in relation to each other and to the outside world; *frontalité*, regulating the distinction between front and back, and the associated functions; and *latéralité*, regulating the arrangement of functions both inside the dwelling and in the farm as a whole to the right and left of the 'master' as he stands at the front entrance of his dwelling welcoming guests. The concept of latéralité is of particular interest in a spatial analysis of the domestic interior, since it specifies not only a principle for the arrangement of rooms, but also a male-centred view of this arrangement. It will be of interest to see how far systematic analysis supports this spatial concept and its social interpretation.

Procedure of analysis

The initial study reported here was based on the largest of the regional samples provided by the Musée: that from Normandy, consisting of 17 dwellings.[13] A preliminary review of this – strikingly heterogeneous – material suggested that in view of the predominance of farm-related dwellings, it would be useful to distinguish three possible levels of analysis:

1. the level of the *minimum living complex*, defining this as the least continuous interior set of spaces which linked together the main living spaces, plus whatever functions formed part of that complex;

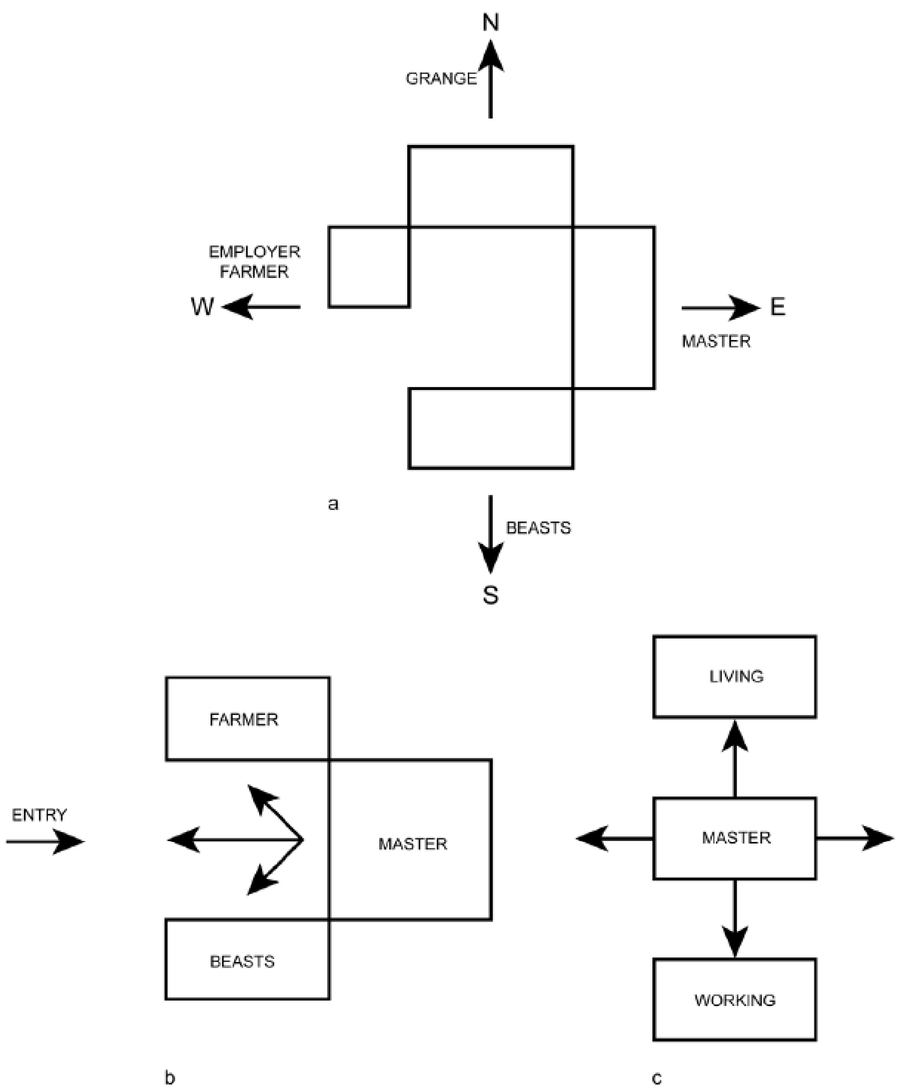


Figure 11.5 Three elements of a model of a farmhouse: (a) *orientation*; (b) *frontalité*; (c) *latéralité*.

2. the minimum living complex plus a single space representing the exterior of the dwelling;
3. the whole complex including outbuildings and spaces only accessible from the minimum living complex through the exterior space.

The analysis reported here deals only with the minimum living complex, with and without its exterior. The complex with its exterior is always dealt with in the first instance. When the exterior is 'discounted' in the analysis, the reader may assume that the syntactic analysis has been reworked and re-calculated for the interior complex only.

The procedure adopted was as follows:

(a) First, justified permeability graphs were drawn for the minimum living complexes, using the exterior as root, whether or not this could be seen as a single space (because we are interested first and foremost in the interior pattern of space, including its relations to the outside, but not, at this stage, in the external differentiation of space per se). These graphs, with their respective plans, are given in Figure 11.6, in the order in which they were originally presented to us.
(b) Second, syntactic analyses of the spatial patterns were made without considering the labels or functions assigned to particular spaces. The resulting data are tabulated in Table 11.1.
(c) Third, the spatial patterns were analysed in terms of functions to see how different functions fitted into the spatial pattern as a whole. The data from this analysis are tabulated in Table 11.2.

Each of these stages of analysis can be expected to generate 'geographical' statements about the sample as a whole as well as 'phenotypical' statements about individual dwellings. The presentation of the analysis will be divided accordingly. First, each house will be commented on as an individual case, drawing on all three types of analysis. Then the sample as a whole will be reviewed, again using all three types of data.

House-by-house analysis

For the house-by-house analysis, the reader should refer in the first instance to the plans and justified graphs in Figure 11.6. This figure also provides a key to room functions. Other material, mainly numerical, will be drawn from Tables 11.1 and 11.2.

House 1 appears at first sight to be a simple linear plan, but the justified graph shows a good deal of morphological differentiation among the spaces. Three spaces in the minimum living space are at depth 1, that is, linked directly to the exterior. Of these, two are transition spaces and one is a function space: the *salle commune*. The salle commune also has a property which is not at all clear from the plan, but made clear by the justified graph: it lies on all three non-trivial circulation rings (that is, those involving more than two spaces). Of these three rings, two are external (that is, pass through the exterior) and one is internal. The internal ring passes through several work-related spaces, including the *laiterie* and the *laverie*. Of the two external rings, one simply links the salle

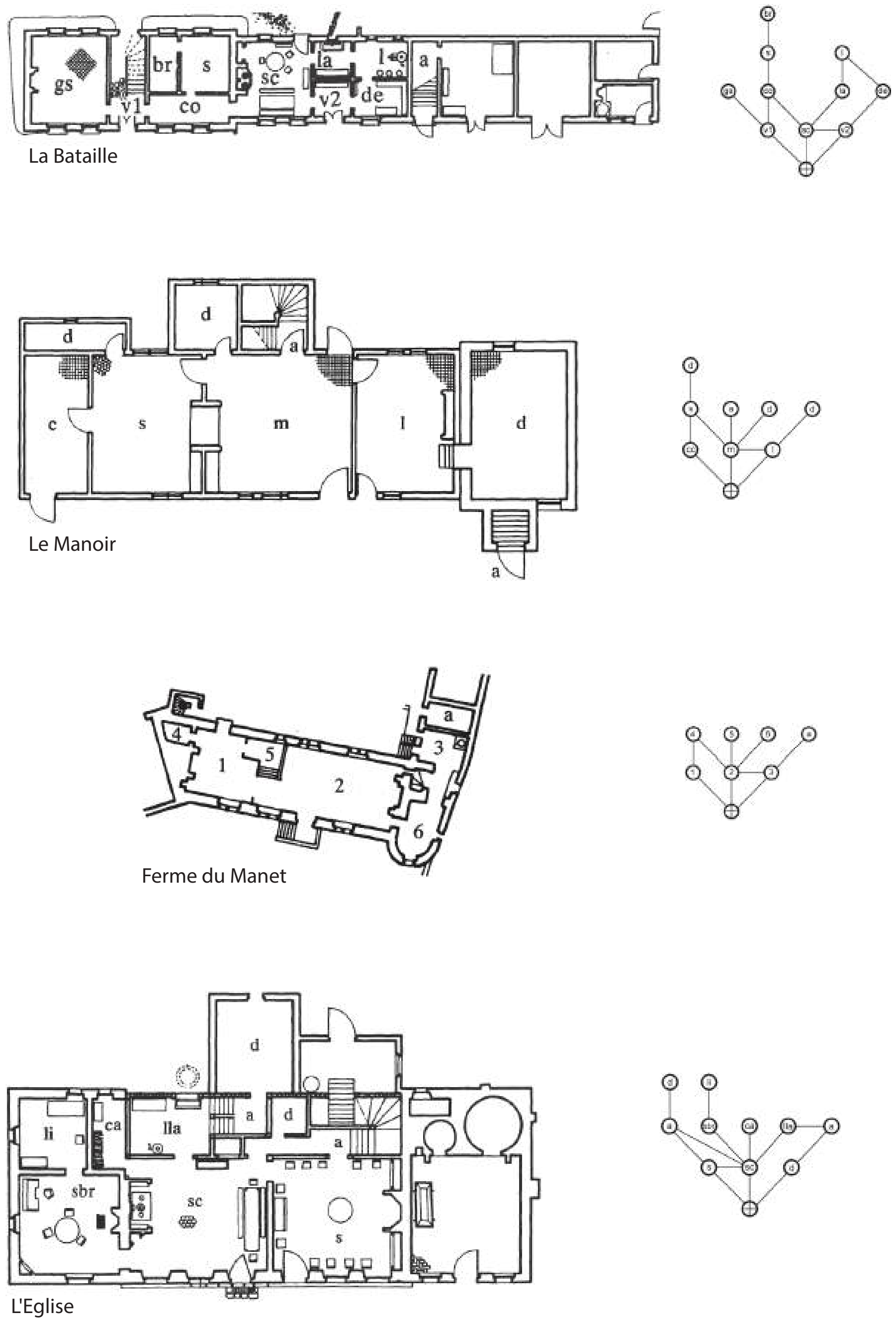


Figure 11.6 Plans of the houses studied and their corresponding permeability graphs. The rooms that are not labelled are storerooms, barns, cowsheds and so on.

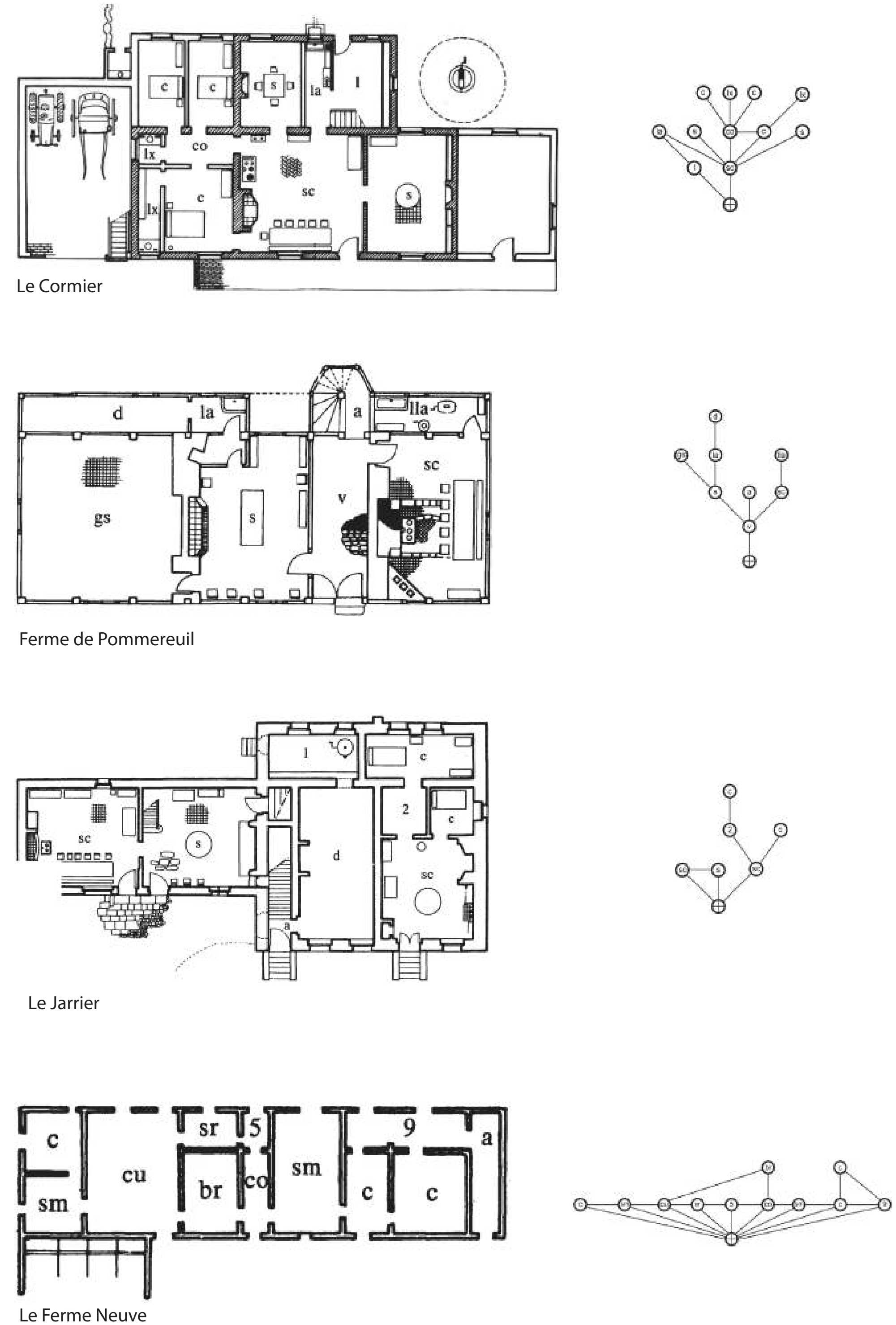
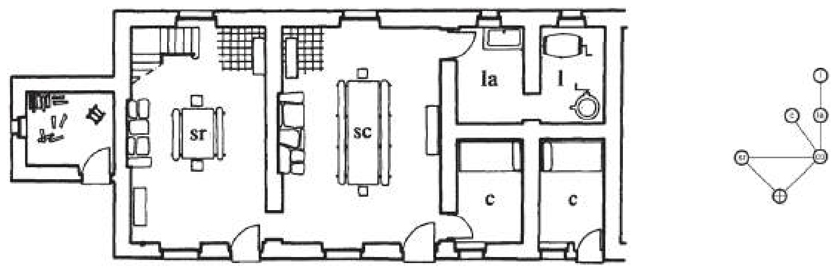
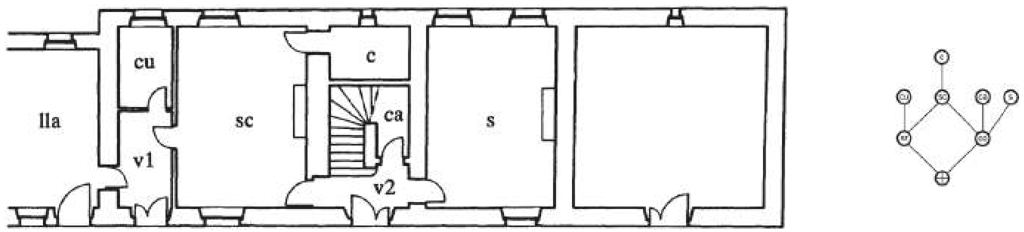


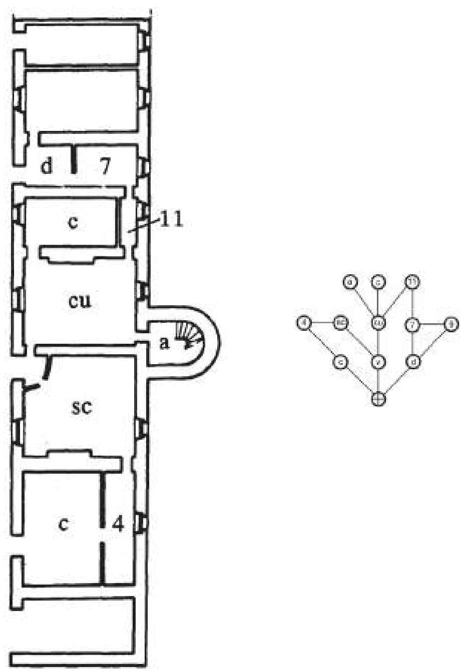
Figure 11.6 (continued).



Le Marais

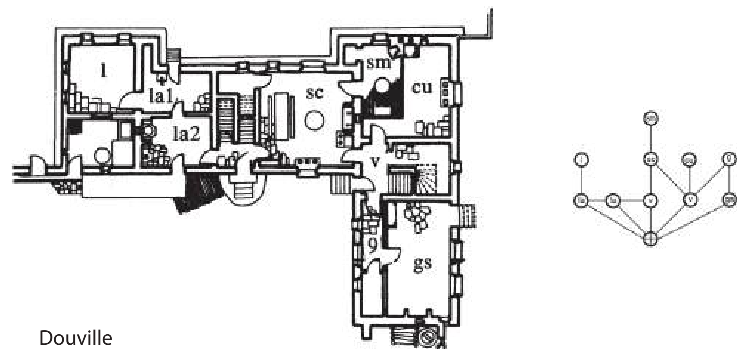


Dodainville, Les Gossets

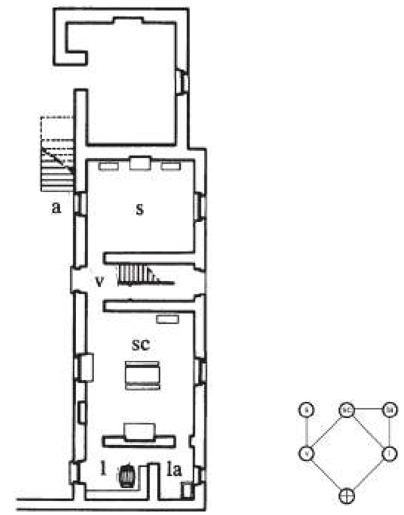


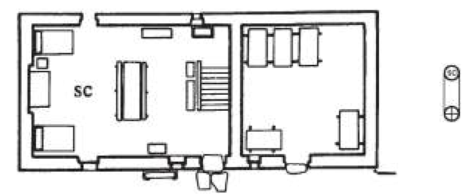
Le Quesnay De Bas

Figure 11.6 (continued).



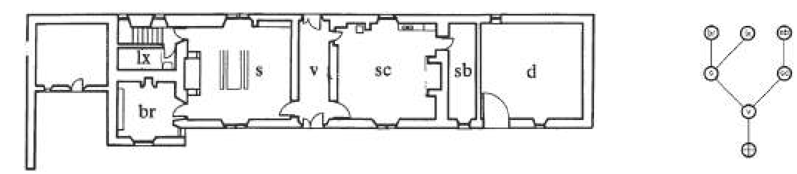
Douville





La BazoQue, Au Village

La Domaine



La Tourps

Figure 11.6 (continued).

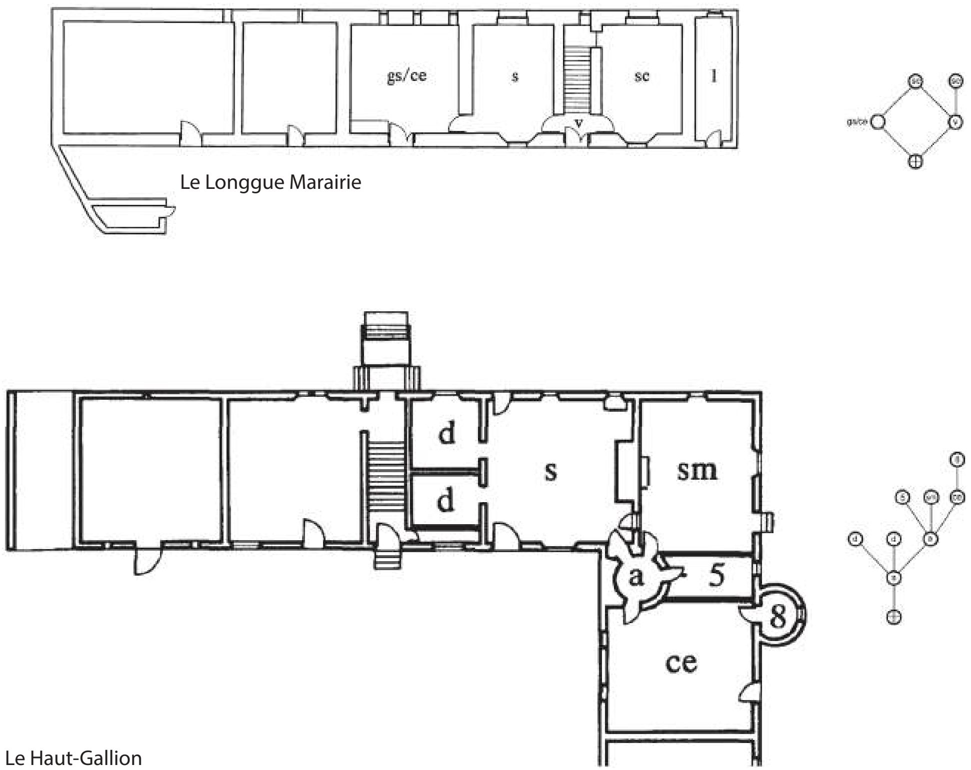


Figure 11.6 (continued).

commune to the exterior by way of a vestibule, but the other is the main link from the salle commune to the other living functions, including the salle, the *grande salle* and the *bureau*. The salle commune in effect acts as a kind of hinge linking and separating two functionally differentiated circulation rings.

Table 11.2, which sets out the integration values of all the spaces in order of integration, shows that the salle commune is also the most integrating space, and by far the most integrating of the function spaces. A strong *inequality* thus exists among the main living spaces with the order: salle commune < exterior < salle < grande salle (meaning that the salle commune is more integrating than the exterior which is more integrating than the salle which is more integrating than the grande salle). All this remains the case when the exterior is discounted, although in this case the *couloir* takes on an equal value to the salle commune as the most integrating space.

The salle commune thus has a striking set of syntactic properties: it is the most integrating space; it lies on all rings; it is shallow from the exterior; and it links and separates the two main functionally differentiated zones of the house.

Table 11.1 Basic syntactic data

House number	Number of cells	Space–link ratio[a]	Integration with exterior			Base difference factor[b]	Integration without exterior			Base difference factor[b]
			mean	min.	max.		mean	min.	max.	
1	10	1.36	1.12	0.60	1.88	0.76	1.36	0.73	2.00	0.81
2	8	1.33	0.95	0.34	1.47	0.66	1.23	0.44	1.80	0.68
3	7	1.25	1.02	0.58	1.45	0.84	1.74	0.98	2.75	0.80
4	10	1.27	0.93	0.30	1.51	0.60	1.22	0.45	2.09	0.62
5	11	1.17	0.89	0.31	1.34	0.66	0.97	0.37	1.58	0.66
6	8	1.00	1.30	0.54	2.03	0.71	1.45	0.73	2.18	0.79
7a	2	1.33								
7b	4	1.00	1.52	0.47	2.37	0.61	2.00	1.00	3.00	0.78
8	11	1.02	0.60	0.13	1.09	0.42	1.52	0.90	2.41	0.82
9	5	1.17	1.15	0.29	2.00	0.49	1.52	0.47	2.37	0.61
10	7	1.13	1.12	0.58	1.74	0.78	1.40	0.59	2.16	0.72
11	11	1.17	1.10	0.67	1.60	0.86	1.71	0.90	2.71	0.78
12	10	1.36	0.96	0.45	1.68	0.69	1.67	0.91	2.73	0.78
13	5	1.50	0.96	0.57	1.72	0.76	1.33	0.47	2.37	0.59
14	1	1.50								
15	6	1.14	1.40	0.59	2.16	0.72	1.62	0.86	2.58	0.78
16	4	1.20	1.14	0.47	1.89	0.68	2.00	1.00	3.00	0.78
17	8	1.00	1.15	0.45	1.80	0.68	1.23	0.44	1.89	0.66
Mean		**1.25**	**1.08**	**0.46**	**1.73**	**0.68**	**1.50**	**0.70**	**2.35**	**0.73**

[a] The space-link ratio is the number of links plus one, over the number of spaces. A tree will therefore have a value of 1, and values above 1 indicate the degree of 'ringiness' in the complex.
[b] The base difference factor is the difference factor for the mean, minimum and maximum integration values in the complex, and thus gives some indication of how much differentiation is available in that complex, which may, or may not, be taken up by the various functions.

It will be of interest to see how far these four properties are reproduced in other cases.

The three other main living spaces – the salle, the grande salle and the bureau – all have quite different syntactic characteristics. All three are non-ring spaces, being either endpoints or on the way to endpoints (see cited texts for a theoretical discussion of this property). The bureau is both an endpoint and also the 'deepest space' in the complex. It is also the most segregated space if the exterior is included, and equally the most segregated if the exterior is discounted. The grande salle is also an endpoint, and the second most segregated space in the complex, if the exterior is included, and equally the most segregated space with the bureau,

Table 11.2 Order of integration of functions, house by house

House Number	Order[a]										
1	SC <	co <	ex =	V <	V <	la <	s <	d <	I <	gs <	b
	0.60	0.68	0.83	0.83	0.90	1.06	1.21	1.28	1.51	1.58	1.88
2	m <	ex =	s =	1 <	c <	d =	a <	d =	d		
	0.34	0.68	0.68	0.68	1.01	1.13	1.13	1.47	1.47		
3	ex =	2 =	1 <	3 =	6 <	a =	5 =	4			
	0.58	0.58	0.58	1.01	1.01	1.45	1.45	1.45			
4	sc <	ex =	s <	lla =	a <	sbr <	ca <	d <	a <	d <	1
	0.30	0.68	0.68	0.75	0.75	0.82	0.98	1.13	1.21	1.43	1.51
5	sc <	co <	c <	ex =	la <	s =	s <	c =	l =	c <	lx <	l
	0.31	0.45	0.57	0.83	0.83	0.96	0.96	1.08	1.08	1.08	1.21	1.34
6	v <	s <	sc <	la <	ex =	a <	gs <	Ila <	d		
	0.56	0.68	1.13	1.24	1.35	1.35	1.47	1.92	2.03		
7b	sc <	2 <	ex =	c <	c						
	0.47	0.95	1.89	1.89	2.37						
8	ex <	c =	sm =	co =	cu <	9 =	5 =	sr =	sm <	c <	br <	c
	0.13	0.51	0.51	0.51	0.51	0.57	0.57	0.57	0.57	0.70	0.89	1.09
9	sc <	la <	ex =	sr <	c <	l					
	0.29	0.86	1.15	1.15	1.43	2.00					
10	v =	SC <	ex =	v <	s =	ca =	l <	cu			
	0.58	0.58	0.87	0.87	1.45	1.45	1.45	1.74			
11	v <	ex <	cu <	d <	11 <	7 =	sc <	c <	4 <	c =	a <	re
	0.67	0.70	0.77	0.96	1.02	1.09	1.09	1.15	1.34	1.40	1.40	1.60
12	ex <	v <	la =	v <	sc <	la <	gs <	9 <	cu <	sm <	l
	0.45	0.53	0.75	0.75	0.83	0.90	0.98	1.06	1.21	1.51	1.58
13	sc =	v <	l =	ex <	la <	s					
	0.57	0.57	0.86	0.86	1.15	1.72					
15	v <	s <	sc <	ex <	la =	br <	sb				
	0.59	0.79	1.18	1.59	1.77	1.77	2.16				
16	v <	ex =	s <	gr <	sc						
	0.47	0.95	0.95	1.42	1.89						
17	a <	s <	ce <	sm =	5 <	ex =	d =	d <	8		
	0.45	0.56	1.01	1.24	1.24	1.35	1.35	1.35	1.8		

[a] For key to room functions, see Figure 11.6; ex = exterior.

if the exterior is discounted. Unlike the salle and the bureau, however, the grande salle is also shallow in the complex. The salle is a relatively deep and relatively segregated space, both with and without exterior, but less segregated than either the bureau or the grande salle.

Of the remaining spaces, all the work-related spaces – laverie, laiterie and *débarras* – are segregated, but all less so than either the grande salle or bureau. If the exterior is discounted, the laverie is a little less segregated. Among the transition spaces, the couloir is a strong integrator, with and without exterior, but the vestibules much less so.

It is also useful to look at the degree of differentiation among the integration values of the different functions. The three main living spaces, for example – the salle commune, the salle, and the grande salle – have a mean integration value of 1.13, but a difference factor of 0.83, which indicates a strong degree of differentiation among the values. In fact, this differentiation among the living spaces is almost as great as it could be in that complex, because, unusually, the most and least integrating spaces are living spaces. If the bureau is substituted for the salle, then the difference factor is even stronger at 0.77.

If, on the other hand, we take the three main work-related spaces – the laverie, the laiterie and the débarras – then the mean integration, at 1.28, is only a little higher than for the living spaces. But the difference factor for these spaces is very weak at 0.97. For the three transition spaces – the couloir and the two vestibules – the mean integration is 0.8, but the difference factor at 0.98 is even weaker than for the work-related spaces. Both of these difference factors are weaker than the ones we obtain by taking the mean integration values of the three types of space (living, working and transition): 0.95, even though this averages out the differences between individual spaces.

These difference factor results are striking and unusual. It is not common to find such strong differences between living spaces, nor for these functions to take up so much of the possible differences in a spatial complex. It will be of interest to see how far the strength and the order of these differences are reproduced elsewhere in the sample.

House 2 has a justified graph which has certain striking resemblances to house 1. Most notably, there is a space which has all four syntactic properties of the salle commune: it is the most integrating space; it is shallow; it lies on all rings (although there are only two, and both external); and it links and separates living from work functions. In this case, however, the space is labelled *maison*, but there seems to be strong functional and syntactic grounds for regarding this space as similar to the salle commune.[14]

There are also a number of differences: there is no grande salle; the salle links directly rather than indirectly to the most integrating space; the work-related spaces to the right of the maison form a dead-end sequence rather than a ring sequence; there are no transition spaces and no bureau; and there is a *chambre de commis* (bedroom for a clerk) on the ground floor. The plan is thus, in certain respects, less spatially complex and less functionally differentiated than house 1. Nevertheless, the justified map shows a striking syntactic resemblance.

This resemblance is reinforced by numerical analysis. The order of integration of the living spaces is maison < exterior < salle < *chambre*, and the difference factor for the three living spaces is again strong at 0.79, much stronger than for the three work-related spaces at 0.89. This, still fairly strong value, reflects the fact that the laiterie is a relatively integrating space in the complex (though much less so if the exterior is discounted).

House 3 is a plan without function data, and geometrically dissimilar to the two previous cases. But the justified graph does suggest certain resemblances. Most striking is that there is a space – marked 2 on the plan – which has the three spatial properties of the salle commune or maison: it is shallow, it lies on all rings, and is the most integrating space in the complex. (This is, however, only the case if one ignores two small spaces – seemingly too small to count as rooms – attached to space 1. If either is included, then it equalises the integration values of spaces 1 and 2.) On the other hand, the external space in this instance is divided into front garden, inner courtyard and approach road (leading to the side door), and cannot realistically be treated as a single space. If this is corrected, then space 2 does become the most integrating space. There is also a comparable difference factor of 0.78 for the three main spaces, and a comparable mean integration for the whole complex of 1.02. However, if space 2 is a salle commune, it is unclear how the other spaces are to be functionally interpreted, and it is perhaps safer to note the syntactic resemblances, but not to speculate too far on the assignment of functions.

House 4 has a salle commune with all four defining characteristics noted for house 1, with one internal and two external rings, and a comparable mean integration of 0.93 for the complex. The order of integration for living spaces is, as before, salle commune < exterior < salle < *salon/*bureau, and the difference factor is again strong at 0.82. In this case, however, the salle is both shallow and on an exterior ring, while the *salon-bureau* is deep and on a dead-end sequence. Internal work functions are, again, on an independent ring sequence linked to the exterior.

House 5 again has a salle commune with all four defining characteristics, with one internal and one external ring. It has two salles, both of which are endpoints, and a relatively integrating chambre lying on an external ring. The mean integration is 0.89, and the order of integration is salle commune < chambre < exterior < both salles. Difference factor for salle commune, chambre and salle is very strong at 0.76. The laiterie is the most integrating work function at 0.83.

House 6 is spatially unlike any previous case, although its functional labelling is familiar in that salle commune, salle and grande salle are the main spaces. As the justified graph shows, the spatial form is that of a tree with a single entrance: it has no rings, internal or external. The salle commune is, however, relatively segregated at 1.13, and the salle is the most integrated of the living spaces at 0.68, reversing the previous order. The mean integration of the complex is 1.30, substantially more segregated than previous cases, and the order of integration for living spaces is salle < salle commune < exterior < grande salle. The exterior is substantially more segregated than any previous case at 1.35, and the most integrating space of all is the vestibule at 0.56. The difference factor for salle commune, salle and grande salle is weak at 0.89, and only by including the vestibule can strong difference factors be found.

House 7 is a special case, since it is split into two distinct living complexes, each with its own salle commune. The one on the left is very simple: a salle commune and a salle connected directly to each other and to the outside, meaning that is maximally shallow, maximally integrating and minimally differentiated. Little can be said of typological interest except perhaps that the salle commune does preserve the spatial characteristics previously noted of being shallow, integrated, and on all rings, but, obviously, not uniquely so.

The right-hand complex is a simple tree form, with the salle commune shallowest and most integrating, and controlling access to two chambres, one directly and one indirectly. The complex is, as a whole, relatively segregated at 1.52, but the salle commune is a strong integrator at 0.47, giving a very strong difference factor for living spaces of 0.61. In spite of its simplicity, the right-hand complex does reproduce the order of integration: salle commune < exterior < chambres. In spite of their differences, therefore, left-hand and right-hand complexes can both be said to reproduce at least some of the spatial characteristics found in most previous cases.

House 8 is both spatially and functionally quite unlike any house so far. Every space, barring the bureau and one chambre, is directly linked to the outside, creating a complex with nine external and two internal rings. The complex is highly integrated at 0.60 if the exterior is included, and very segregated at 1.52 if it is discounted. (In this case, it is unrealistic to treat the exterior as a single space, since it divides sharply into rear walled garden, inner courtyard and outside proper. However, even with the garden treated as a separate space, the mean integration with the exterior is 0.64 and the complex behaves in a very similar way.) There is no salle commune, no salle and no grande salle, but there is a bureau which combines the properties of being one of the two spaces 2-deep in the complex, of being the second most segregated function space (one chambre is more segregated) if the exterior is included, but the most integrated function space if the exterior is discounted.

Instead of the more common function spaces, there is a *cuisine*, a *salle à manger des maîtres* – and two other salles à manger, one directly linked to the cuisine and small, the other large, separated from the cuisine by two intervening spaces, and said to be '*d'apparat*', meaning that it is for special occasions. Difference factors for the major spaces are very weak: 1.00 for large salle à manger, cuisine and bureau without the exterior, 0.91 with the exterior. Without the exterior, the couloir is the most integrating space, and with the exterior it is equally most integrating with the cuisine and the large salle à manger. The order of integration changes with, and without, the exterior, and either way is unlike any previous case. With the exterior we find exterior < couloir = large salle à manger = cuisine = small chambres < bureau < large chambre. Without the exterior, we find couloir < bureau < large salle à manger < cuisine < chambres. Both with and without the exterior, the salle à manger des maîtres is average in integration, but with the bureau, it seems to divide the house into two zones. Both functionally and spatially the division suggests a fundamental distinction between masters and servants rather than between living and working.

House 9 has a much simpler salle plan, and returns, in a simplified form, to some of the salle commune themes. The salle commune is the most integrating space, lies on the only (external) ring, is shallow and separates living from work functions. This time, however, the salle commune is described as '*des domestiques*'. There is neither salle, nor grande salle, but there is a salle des maîtres, and this is much larger than in the previous case. Spatially, the salle des maîtres seems comparable in some respects with the salle, in that it is less integrated than the salle commune, but more integrated than the chambre. On the other hand, it is both shallow and lying on the external ring, and in this resembles the normal salle commune. Mean integration for the complex is normal at 1.14, and difference factors are strong with 0.62 for salle commune, salle des maîtres and chambre. The order of integration is: salle commune < exterior < salle des maîtres < chambre. The bureau is external and independent and does not form part of the minimum living complex.

House 10 introduces some new features into a pattern that nevertheless continues to resemble the salle commune type. The first, deep in the plan, is a small cuisine, which has not so far co-existed with a salle commune. The second is a pair of vestibules – one resulting from the same partitioning that created the cuisine – which unlink the salle commune from the exterior. Even so, at 0.58, the salle commune remains the most integrating function space, equal to the central vestibule, when the exterior is included, and easily the most integrating space at 0.59 if the exterior is discounted. The cuisine is the most segregated space both with (1.74) and without (2.16) the exterior. The salle is also strongly segregated. Mean integration is average at 1.12, and the order of integration for living spaces is salle commune < exterior < chambre < salle < cuisine. Difference factors are strong with 0.83 for salle commune, salle and chambre, and 0.79 for salle commune, salle and cuisine.

House 11 is another rare case where a cuisine co-exists with a salle commune, though in this case the cuisine has become, with or without the exterior, the most integrated function space at 0.77, compared with 1.09 for the salle commune. With the exterior, the rudimentary vestibule is the most integrating space of all, though without the exterior the cuisine takes over. Spatially, the complex is characterised by two deep, external rings, but no internal rings. Mean integration is average, at 1.10, but becomes highly segregated, at 1.71, without the exterior. The order of integration for living spaces is exterior < cuisine < salle commune < chambre. Difference factors are weak with 0.92 for cuisine, salle commune and large chambre.

House 12 has more functional differentiation of living spaces than any other case, with salle commune, cuisine, salle à manger and grande salle. Even so, it reproduces some – but not all – of the features of the dominant salle commune type. With the exterior, the salle commune remains the most integrating living space, but the exterior is much more integrating as are both the central vestibule and one laiterie (because of the strong integration of the exterior). Discounting the exterior, the salle commune becomes uniquely the most integrating space. The order of integration of living spaces with exterior is salle commune < grande salle < cuisine < salle à manger; and, without, salle commune < salle à manger < cuisine < grande salle. Mean integration is normal at 0.96, but this is largely because of the effect of the exterior. Without the exterior, mean integration is 1.67. Difference factors for living spaces are weak with the exterior, with 0.92 for salle commune, grande salle and salle à manger, but become stronger when the exterior is discounted, with 0.88 for the same three spaces. With the exterior, strong difference factors only arise if the central vestibule is one of the spaces considered. Finally, all four rings in this complex are external, but the salle commune does link and separate living and work functions.

House 13 is another case of the dominant salle commune type in simplified form. The salle commune, in spite of being unlinked from the outside by a vestibule, and the laverie, is the most integrating function space (equal to the vestibule at 0.57) with the exterior and by far the most integrating space of all at 0.47 without the exterior. It also lies on both rings, one internal, one external, and links and separates living from internal work functions. Mean integration is average at 0.96, going up to 1.33 without the exterior. The order of integration is salle commune < exterior < salle, following the dominant pattern. Difference factors are strong with 0.78 for salle commune, salle and laverie, but there is not enough living space to compute this for living spaces alone.

House 14 has a single space minimum living complex, and cannot therefore be analysed. Even so, the fact that the single space – which must already be shallow and integrating – lies on a ring and functions as a salle commune is not without typological relevance. It could be argued that the complex would only have to develop to preserve the features that are already present, to arrive at the dominant salle commune type.

House 15 has a salle commune, but clearly does not conform to the dominant type. Spatially, the complex is split by the entrance vestibule into two branches of a tree, with the salle commune on one branch and the salle on the other. Because there is one extra space on the salle side, the salle appears as the most integrating function space, though with the rather poor value of 0.79, compared with 1.18 for the salle commune. The vestibule is the most integrating space, and the exterior is strongly segregated at 1.57. Mean integration is 1.42, with the exterior, and 1.62 without. The bureau is, again, strongly segregated at 1.77 and deep in the complex. The order of integration for living spaces is salle < salle commune < exterior < bureau. Structure factors are fair with 0.87 for salle, salle commune and bureau, but this is more a result of the segregation of the bureau than of the strong integration of any spaces. It is perhaps worth noting that several of the properties of the dominant salle commune type would be restored if the – apparently added – partition between the *salle de bains* and the débarras were removed.

House 16, although spatially it could approximate a simplified version of the dominant salle commune type, in fact inverts it by having the salle commune as the most segregated space at 1.89 and the only endpoint. The salle both integrates more than the salle commune and lies on the single exterior ring, but it integrates less than the vestibule. Mean integration is normal at 1.14 with the exterior, but if the exterior is removed, the complex becomes a single sequence of spaces with a mean integration of 2.00. Difference factors for function spaces are very weak, in spite of the strongly segregated salle commune, but become very strong if the vestibule is considered as one of the spaces – for example, salle commune, salle and vestibule have 0.65.

House 17 is another tree form, without a salle commune, but with a salle as the most integrating function space at 0.56 and a deep transition space as the most integrating space of all at 0.45. Mean integration is average, at 1.15 with the exterior, and 1.23 without, showing that integration depends little on the exterior. The order of integration for living spaces is salle < salle à manger < exterior. Difference factors are weak unless the transition space is taken into account, in which case we find 0.75 for transition space, salle and salle à manger.

The problem of type

The house-by-house review suggests that, although there is no obvious single house 'type' in the sample – defined perhaps as a more or less standard way of constructing the house and arranging its rooms – there is evidence of at least one underlying spatial-functional 'genotype' – defined in terms of relational and configurational consistencies which show themselves under different 'phenotypical' arrangements.

However, sometimes this dominant genotype is realised strongly, in that all the spatial-functional themes are present, sometimes more weakly, in that some are present and some are missing, whereas in other cases these themes seem to be totally lacking, or even inverted.

The questions to be addressed in this section therefore are: 'can the idea of a dominant genotype be formally demonstrated?' and, 'is there a second type, and can this be formally demonstrated?' The first step in trying to answer the first question is to consider the spatial and functional properties of the sample as a whole. Table 11.3 sets out each main type of space that occurs in the sample, the number of times it occurs, and its mean depth and integration value when it does occur. This shows that the commonest types of function space are salles communes and chambres, with 13 each, then salles, followed by transitions and various work spaces. Cuisines are rare, as are grandes salles.

There are also clear, across-the-board differences in the way in which these functions are spatialised. Salles communes occur in the sample with a mean depth of 1.47 and a mean integration value of 0.74 (0.79 without the exterior); salles with a mean depth of 1.91 and a mean integration of 1.01 (1. 13 without the exterior); grandes salles with a mean depth of 2.00 and a mean integration of 1.34 (2.00 without the exterior); and chambres with a mean depth of 2.07 and a mean integration of 1.21 (1.67 without the exterior). These differences are sufficient to give a difference factor of 0.93 for these means for salle commune, salle and grande salle, which would not be strong in an individual case, but is strong in a sample.

Among the less common spaces, *cuisines* are rare, but where they occur their mean depth is 1.75 and mean integration 1.06 (1.52 without the exterior). *Cuisines*, in effect, only appear occasionally and in deep and segregated spaces. *Salles à manger* are similar, but the two salles des maîtres are both shallow and relatively integrating. *Bureaux*, on the other hand, are, on average, strongly segregated at 1.34. Work functions are in general considerably more segregated than living functions, and there are fewer differences among them. Laveries are both the deepest of all function spaces and the most integrating of the work functions at 1.15. Transition spaces, on the other hand, are common, and on average both shallow and strongly integrating. The overall mean integration for all spaces in the sample is 1.08, and very broadly one might say that living functions are on the integrated side of the mean and work functions on the segregated side.

Table 11.3 Numbers, mean depths and mean integration values for functions

		With exterior		Without exterior
Function	**Number of cases**	**mean depth**	**mean integration**	**mean integration**
Exterior	16		0.93	
Salle commune	13	1.47	0.74	0.79
Chambre	13	2.07	1.21	1.67
Salle	11	1.91	1.01	1.13
Vestibule	9	1.00	0.68	0.95
Laverie	9	2.20	1.15	1.42
Laiterie	8	2.00	1.33	1.76
Cuisine	4	1.75	1.06	1.52
Salle à manger	4	2.00	0.96	1.45
Grande salle	3	2.00	1.34	2.00

These strong trends across the sample are, in themselves, strong evidence of an underlying spatial culture expressing itself through the spatial form of the houses. However, this spatial culture expresses itself in spite of the numerous inversions and oppositions that were noted in the house-by-house review. It seems likely, then, that if more than one genotype could be identified, spatial cultures would show through and be expressed even more strongly.

A commonsense, conjecture-test procedure seems most appropriate. The house-by-house review suggested a dominant type based on the existence of a salle commune with the four properties of being shallow, most integrating, lying on all rings, and linking and separating living from work functions. House 1 seems a clear case; house 2 can be allowed since maison and salle commune are used interchangeably elsewhere; house 3 is unlabelled and must be omitted; but house 4 is clear, as is house 5. House 6 is clearly not a case, and house 7 is a reasonable case, but perhaps should be omitted as being too small. House 8 is not a case, but houses 9 and 10 reasonably are. House 11 is not a case, but house 12 is. House 14 is too small, and then houses 15–17 are all, clearly, not cases. Disregarding the houses which are too small, we thus have eight possible cases of the dominant genotype and six cases which clearly do not conform to this genotype.

Table 11.4 divides the sample into two along these lines, showing mean integration with and without the exterior, the function and integration value of the most integrating space, the difference factor for the main living spaces, and the integration value for the exterior. The table shows a number of interesting results:

1. The mean integration of the genotype sample is very stable at around 1. The houses that strongly deviate from the mean are all in the nongenotype sample, which also has a slightly higher mean.
2. If the exterior is discounted, the mean integration of the genotype sample is much stronger at 1.37 than the nongenotype sample at 1.59.
3. The mean integration for salles communes in the genotype sample is 0.47 with the exterior and 0.6 without; for the nongenotype sample the mean is 1.32 with the exterior and 1.81 without.
4. The salle commune is the most integrating space of all throughout the genotype sample. The only exceptions are house 10, where the vestibule is equally most integrating if the exterior is included, but the salle commune is uniquely and strongly most integrating if the exterior is discounted; and house 12, where the exterior is the most integrating space and the salle commune only the most integrating living space if the exterior is included; but again the salle commune becomes uniquely and strongly most integrating if the exterior is discounted.
5. A quite different, yet consistent, pattern of most integrating spaces is found in the nongenotype sample: in house 6, the vestibule is most integrating, although the salle is equally so if the exterior is discounted; in house 8, the exterior is by far the most integrating space, but the couloir follows, and becomes most integrating if the exterior is discounted; in house 11, the vestibule is the

Table 11.4 Data on two possible types of house

House number	Mean integration[a]		Most integrating space[b]				Difference factor for main function spaces	Integration value of exterior
	with	without	with		without			
Genotype								
1	1.12	1.36	sc	0.60	sc	0.79	0.83 (sc, s, gs)	0.83
2	0.95	1.23	sc	0.34	sc	0.44	0.79 (m, s, c)	0.68
4	0.93	1.22	sc	0.30	sc	0.45	0.82 (sc, s, sb}	0.68
5	0.89	0.97	sc	0.31	sc	0.37	0.76 (sc, s, c)	0.83
9	1.10	1.52	sc	0.29	sc	0.47	0.62 (sc, sm, c)	1.15
10	1.12	1.40	sc	0.58	sc	0.59	0.83 (sc, s, c)	0.87
12	0.96	1.67	ex	0.45	sc	0.91	0.88 (sc, gs, sm)	0.45
			(sc	0.83)			(0.92 with ex)	
13	0.96	1.33	sc	0.57	sc	0.47	0.78	0.86
Mean	**1.01**	**1.37**	**sc**	**0.48**	**sc**	**0.56**	**0.79**	**0.79**
			v	**0.64**				
				0.80[b]				
Nongenotype								
6	1.30	1.45	v	0.56	s,v	0.73	0.89 (sc, s, gs)	1.35
			(sc	1.13)	(sc	1.31)		
8	0.60	1.52	ex	0.13	co,v	0.90	0.91 (sm, cu, br)	0.13
			(co	0.51}			(1.0 without ex)	
16	1.14	2.00	v	0.47	s,v	1.00	0.91	0.95
			(sc	1.89)	(sc	3.00)		
11	1.10	1.71	v	0.67	cu	0.9	0.92 (sc, c, cu}	0.70
17	1.15	1.23	a, v	0.45	a,v	0.44	0.88 (s, sm, ce)	1.35
Mean	**1.12**	**1.59**	**v**	**0.55**	**v**	**0.84**	**0.90**	**1.01**
			sc	**1.32**	**sc**	**1.81**		
				0.54[b]				

[a] With and without exterior.
[b] All transition spaces.

most integrating, though it becomes second to the cuisine if the exterior is discounted; in house 15, the vestibule is again most integrating, though again the salle has an equal value if the exterior is discounted; in house 16, the vestibule is again most integrating, though again it is joined by the salle if the exterior is discounted; and in house 17, the small interior transition space is most integrating, remaining so when the exterior is discounted.

6. The mean integration for these transition spaces in the nongenotype sample is 0.54 with the exterior and 0.84 without. The comparable figures for transition spaces in the genotype sample are 0.78 with the exterior and 1.02 without – in other words, salles communes and transition spaces change places in the two samples.
7. Difference factors then reflect this change: the mean difference factor for living spaces in the genotype sample is 0.79, whereas for the nongenotype sample it is 0.90. In the nongenotype sample, strong difference factors are only found when transition spaces are included in the space considered, and vice versa in the genotype sample.
8. Last, the mean integration of the exterior in the genotype sample is 0.79, whereas for the nongenotype sample it is 1.01.

In other words, two distinct genotypical tendencies can be demonstrated in the sample. One centres on the highly integrating salle commune, creates strong spatial differences among living spaces, incorporates the exterior in its pattern of strong integration, and has a more integrating interior and a more integrating exterior. The other centres on the transition space, creates more internal segregation amongst living spaces and less spatial differences among them, separates the inside more clearly from the outside, and has a more segregated exterior. These genotypes do not appear to be correlated either with size or with the overall geometry of the building. On the contrary, they appear to be two distinct spatial-functional tendencies, each of which expresses itself through several different built forms.

An interpretative speculation

In considering these two genotypes against the background of the concepts drawn from Cuisenier's interpretation of Estienne,[15] the concept of latéralité, implying the division of the dwelling into living and working zones on either side of a central space, seems particularly apposite. It is a pervasive theme throughout the sample, though with great variation in the way it is realised and the degree to which it is realised.

However, when it is related to the two genotypes, a more complex picture emerges, Cuisenier's model specifies a latéralité with three strong properties: it has a geometric, or left-right element; it is organised around a central transition space; and it is based on the point of view of the male master of the house. None of these properties can be left without further comment.

On the geometric, or left-right question, it is clear that this does sometimes apply, for example in houses 1 or 9. But in other cases, the latéralité is as strongly realised in the syntax of the spaces, but takes on either a front-back geometry, as for example in houses 4 or 5, or a more indeterminate form, as in house 12. It seems more reasonable, on the basis of this evidence, to think of latéralité as a primarily *syntactic* property which sometimes takes one geometric form, sometimes another. It is pervasively present, but its form seems more to do with the cultural arrangement of practicalities than with an exogenous conceptual model.

On the central space question, it is clear that, although latéralité is sometimes organised around a central transition space, more often it is organised around the dominant *function* space: the salle commune. Which alternative is selected seems to be the principal choice that leads to one genotype or the other. This raises an important question: does latéralité organised around a transition space mean the same thing as latéralité organised around a main functional space? Or does it arise in different social circumstances?

This in turn raises the question of the male-centred view of latéralité. The salle commune, with its linking of cooking and everyday living, seems to be a space in which women would be expected to be dominant, the more so since the work functions which the salle commune typically separates from other living functions are those associated with female roles – the laverie, the laiterie and so on. It is difficult to avoid the inference that the salle commune-centred form of latéralité is in fact organised around the female functions of the household. One is almost tempted to the view that the transition-space-centred form of latéralité, following Cuisenier's interpretation of Estienne, is associated with a male view of the household, and the salle commune-centred form with a female view.

However, the attractions of this simple 'explanation' of the two genotypes must, at least, be put in question by an awkward fact: the distinction between transition-space-centred and function-space-centred domestic space organisations has been made before in quite different explanatory circumstances. For example, Hillier and Hanson note such a distinction in distinguishing domestic space styles which express class more than gender differences,[16] whereas Glassie associates such a distinction with social changes over a period of time linked to changes in house locations and changes in privacy needs.[17]

In both of these studies, however, a similar view is taken of the social mechanisms underlying domestic space patterning. Both emphasise the importance of considering the house, not only in terms of the relations among its inhabitants, but also in terms of the relations between inhabitants and visitors. Domestic space cannot be understood without understanding the dynamics of both types of relationship, and the house can only be understood as a device for managing both types of interface. In both studies, the house is thus seen as a spatial and symbolic means to social and communal solidarities, as much as an instrument of family and individual privacy.

In pursuing these ideas, we explore what we might call the experiential dimensions of space, and in particular, the changing experience of the house as one moves from one space to another. A key aspect of this is often the relationship between permeability and visibility. The permeability structure of a complex is essentially a matter of how the relations of spaces to their immediate neighbours builds into a system of possible routes. It defines where you can go and how to get there. The visibility structure, on the other hand, tells you how much space you are aware of without moving. In a sense, it tells you where you already are.

The relations with visibility are often, it seems, a means by which the basic permeability syntax of a complex is fine-tuned into a more effective device for interfacing or distancing different kinds of relationships. This certainly seems true of the Normandy sample. If, for example, one looks at the salle commune in house 1 (assuming doors are open), there is a line of sight and direct access that crosses the salle commune, passes through the couloir controlling access to the salle and bureau, then through the front-back transition space, and then through the grande salle. Another such line crosses the salle commune then passes through the main entrance vestibule to the outside. Another crosses the salle commune and passes through both laverie and laiterie. In a sense, all the major spatial relations in the complex are governed visibly from the salle commune: the interface between the salle commune (that is, space of everyday living) and the other living functions of the house; the interface between the salle commune and the interior work functions; and the interface between the salle commune and the world outside.

In total contrast, in house 11, the visibility relations from the salle commune are hardly more than the immediate neighbouring permeabilities, and even these are highly restricted. None of the three interfaces of visibility that are so evident in house 1 are realised to any degree in this case except, dubiously, that with the outside world. To be in that space is only to be in that space, not to be visibly part of a complex system of spaces, involving both interior and exterior. Similar differences are found if one compares, for example, house 5 with house 6.

In contrast, the most striking cases of visual relationships in the transition space type occur with the transition space itself. Houses 6, 11, 15 and 16, for example, all have the strongest visual relations from the vestibule just inside the main entrance, whereas house 8 has a seven-space enfilade with this point in the couloir as its centre. House 17 does not have this property, but even there, in a less strong sense, the interior transition space is the strongest visual integrator.

These distinctions are, it seems, reinforced by the ring structure. In the salle commune type, the eight salles communes lie on a total of 15 rings, or 1.87 per salle commune. In fact, with the exception of house 12, where the salle commune lies on only one of three rings, the salles communes lie on all rings in the complexes. On the other hand, if the external rings are cut, then, in each case, the salle commune becomes a controlling space which must be passed through to move from one part of the house to another. In contrast, of the four salles communes in the transition space type, only one lies on a ring, and that a single ring. In this type, the transition space becomes the controlling space which must be passed through to move from one part of the house to the other, with much more restricted opportunities to use the exterior for alternative routes.

The salle commune in the salle commune type is, it seems, a controlling space for the interior – its control of certain aspects of interior permeability is unavoidable – but only a strategic space for the interior–exterior relation; it is powerful, but avoidable. The transition space in the transition space type is, on the other hand, more often a controlling space both for interior and for interior–exterior relations.

It is hard to avoid the inference that these relations are linked to the ways in which domestic space creates and structures the possibility and form of encounter among inhabitants and between inhabitants and visitors, and that the differences between the two genotypes express some difference in the forms of social solidarities. The salle commune type seems to suggest a pattern that works by creating spatial differences between functions, strong interior integration with everyday living as the centre, and a permissive rather than controlling relation to the outside world. The transition space type works by more uniformly segregating interior functions through a central transition space which controls both interior relations and relations with the outside.

The first might be seen as a *constitutive* or *spatial* model in which the social role of space is expressed directly through the way in which the space pattern is lived; whereas the second might be seen more as a *representative* or *conceptual* model, in which individual function spaces are assigned a spatial identity more through separation and control than through the organisation of complex interrelations.

Such a distinction may, however, itself be related to the different ways in which gender relations can express themselves through space. The suggestion has been made before.[18] There seems, perhaps, a possibility that we may be dealing with a pair of 'genotypical' tendencies of some generality. But their further exploration would require 'nonarchaeological' forms of data, and thus lies beyond the scope of this present paper.

Notes

1. B. Hillier and J. Hanson, *The social logic of space* (Cambridge: Cambridge University Press, 1984), p. 26.
2. B. Hillier, J. Hanson and H. Graham, 'Ideas are in things: An application of the space syntax method to discovering house genotypes', *Environment and Planning B: planning and design* 14 (1987): 363–385.
3. B. Hillier, *Space is the machine: A configurational theory of architecture* (Cambridge: Cambridge University Press, 1996).
4. B. Hillier et al., 'Ideas are in things', pp. 364–365.
5. B. Hillier and A. Penn, 'Visible colleges: Structure and randomness in the place of discovery', *Science in Context* 4 (1991): 23–49.
6. J. Hanson, *Decoding homes and houses* (Cambridge: Cambridge University Press, 1998), p. 1.
7. B. Hillier, J. Hanson, J. Peponis, J. Hudson and R. Burdett, 'Space syntax: A different urban perspective', *The Architects' Journal* 178(48) (1983): 47–63; P. Steadman, *Architectural morphology* (London: Pion, 1983); Hillier and J. Hanson, *The social logic of space*; J. Peponis, 'The spatial culture of factories', *Human Relations* 38 (1985): 357–390; B. Hillier, 'The nature of the artificial', *Geoforum* 16 (1985): 163–178.

8. This value must be subjected to one more transformation if spaces in graphs with different numbers of spaces are to be compared (Hillier and Hanson, *The social logic of space*, pp. 109–113).
9. Hillier and Hanson, *The social logic of space*, pp. 109–113.
10. C. Shannon and W. Weaver, *The mathematical theory of communication* (Chicago and London: Illinois University Press, 1948).
11. The Unit is the precursor to what is known nowadays as the Space Syntax Laboratory (eds).
12. C. Estienne, *La maison rustique: Logique sociale et composition architecturale* (Paris: J. Du Puis, 1564). Exposition by J. Cuisenier, 1985, 'Type idéal et realités architecturales', draft of Chapter 1 of *La maison rustique*, mimeo.
13. M.-A. Brier and P. Brunet, 'Normandie', in *L'architecture rurale Française: La Normandie*, edited by J. Cuisenier (Paris: Berger-Levrault, 1984).
14. Consultation of the Brier and Brunet text confirms that the two terms are often used interchangeably.
15. Cuisenier, 'Type idéal et realités architecturales', 1985.
16. J. Hanson and B. Hillier, 'Domestic space organisation: Two contemporary space-codes compared', *Architecture and Behaviour* 2 (1982): 5–25; Hillier and Hanson, *The social logic of space*, pp. 151–163.
17. H. Glassie, *Folk housing in middle Virginia: A structural analysis of historic artifacts* (Knoxville: University of Tennessee Press, 1975), pp. 114–122.
18. Hillier and Hanson, *The social logic of space*, pp. 239–240.

12 Against enclosure (1988)

Introduction to 'Against enclosure'

Ann Legeby

It is striking how relevant Bill Hillier's article 'Against enclosure' still is, even in contemporary architectural discussions. The article presents a strong critique of the notions that see social coherence and community to be dependent on a limited, bounded and enclosed geometric form. Hillier demonstrates how housing estate design is in an essential way different from the logic of traditional cities and – in a way – 'Against enclosure' is a story of how fundamental properties of urbanity were abandoned.

The principle of designing housing estates as spatially separated enclosures, clearly identifiable from their surroundings, became a common solution to designing housing in the past century. It was based on the assumption that space can only be socially significant if a specific group of residents was identified with it and the architecture should as far as possible exclude others from the internal spaces of housing units. Anyone not ostensibly belonging to that unit could be identified as an 'outsider', an intruder and, as such, a threat to the local community. Such principles, where different categories are spatially separated (for example locals and non-locals) have, however, been strongly contested, for example by Jane Jacobs, in her influential work on the importance of street diversity to create lively cities.[1]

Hillier claims that the design principles of enclosure, repetition and hierarchy have proven to be a persistent myth. He shows that the principle was already explicit in early twentieth-century designs by Bofil and Le Corbusier, who treated enclosure as being inherently a social good, a principle later advocated by proponents such as Oscar Newman.[2] 'Against enclosure' challenges these ideas and uses space syntax to illustrate the correspondence between urban form and urban life, showing how architecture creates conditions for encounter, co-presence, co-awareness and co-absence.[3] The article shows that paradigms of enclosure-repetition-hierarchy did not provide favourable conditions for encounter. Instead, such design creates fragmented, unintelligible and largely under-used spaces. The article 'Against enclosure' was first presented at the *Rehumanizing Housing Conference* in 1987.[4] Hillier addressed the oversimplified solutions coming from the defensible space thinking, advocated by Newman and Coleman, and instead identified the characteristics of the morphology and the spatial relations (the syntax) and how this corresponds to certain conditions that may increase the vulnerability for social exclusion and antisocial behaviour. In a sense, this represented less a deterministic posture than a probabilistic view.

Hillier's position is substantiated by empirical data based on his and Hanson's meticulous research. This shows that design principles of enclosure result in sparsely populated streets and a non-urban character. So, what are the social effects of such design principles? Contrary to traditional cities – which Hillier describes as mechanisms for generating contact – the fragmented city instead becomes a mechanism of segregation. This, he argues, is due to a reduction in spatial scale, a lack of internal spatial structure within enclosed housing areas, and the resulting concentration of the little movement that there is in the peripheral streets. The design of the estate disconnects it from the larger city context. Taken together, these factors constitute a betrayal of long-standing traditions of designing for social interaction. In fact, such urban environments are on a fundamental level different from traditional (street-based) cities. Hanson later formulated a very eloquent description of the new design principles being in essence a transformation from an all-neighbour design approach to one of no-neighbours, characterised by a ruptured interface between housing and streets, and a layout that minimises social contact.[5] The mismatch between the local and the global scales of street configuration and its implications for social co-presence has been further elaborated within the space syntax field, where it has been shown that the distribution of movement and co-presence is a function of spatial configuration over and above the impact of land uses.[6] Subsequent research has confirmed that urban form has the ability to create potential for people to be co-present in public open space and through that, share their everyday experiences – this outcome being especially important in discussions regarding urban segregation.[7] Similarly, Vaughan and colleagues conclude that suburban town centres have evolved to take advantage of differing scales of movement and encounter over time – here substantiating the argument that there are essential factors that make for the urbanity of cities.[8]

'Against enclosure' still carries a highly relevant message. In contemporary society, we see increasing social polarisation alongside trends toward urban segregation. There is a risk that current urban design practice may reinforce patterns of encounters being limited to in-group contact, rather than across disparate societal groupings. This calls for urban planning and design to create the conditions for streets to be easily shared with others and that enable us to live among strangers as well as provide opportunities for the co-existing of groups that may foster social cohesion in a broader sense than only limited to the local community. Moreover, design practice has every reason to be on its guard; similar ideas pop up regularly, as a wolf in sheep's clothing. What Hillier pointed out in 'Against enclosure' back in the late 1980s is still true: 'the idea of the localised enclosure is not new. Nor is it dead'.

Against enclosure

Bill Hillier

Originally published: Hillier, B. 1988. 'Against enclosure'. In: Teymur, N., Markus, T. and Wooley, T. (eds), *Rehumanizing housing*. London: Butterworth, 63–88.

Introduction: Enclosure, repetition, hierarchy = fragmentation

Architectural ideas typically associate social values with spatial concepts. In the recent past, a common social value in housing design has been that of the small, relatively bounded community, forming an identifiable unit of a larger whole. Architecturally this has been reflected in a preoccupation with linking groups of dwellings to identifiable and distinct external spaces in the hope that the 'enclosures' or 'clusters' so created would help group identification and interaction. The idea is justified spatially by invoking urban squares, courts and village greens, and socially through notions of 'group territory', the 'need for a hierarchy from public to private space', and the assumption that space can only be socially significant if a definite group of people are identified with it.[9]

In its extreme form, enclosure becomes the basis for a methodology of layout design in which local enclosures are either repeated or subjected to simple geometrical transformations, then reproduced at a higher level to create an 'enclosure of enclosures', or a similar hierarchical design. Figure 12.1, a design for a village in Algeria by Ricardo Bofill, is a perfect summary of these three principles of enclosure, repetition and hierarchy, working at three levels. Dwellings are first wrapped round a small, local space. A set of these composite units is then wrapped around a larger space. Then these second-order composites are wrapped around a central 'square'.

Figure 12.2 shows an international selection of schemes from the book *Residential districts*,[10] in which over a hundred housing schemes from many countries are reviewed. In spite of their geometrical dissimilarity, all the schemes

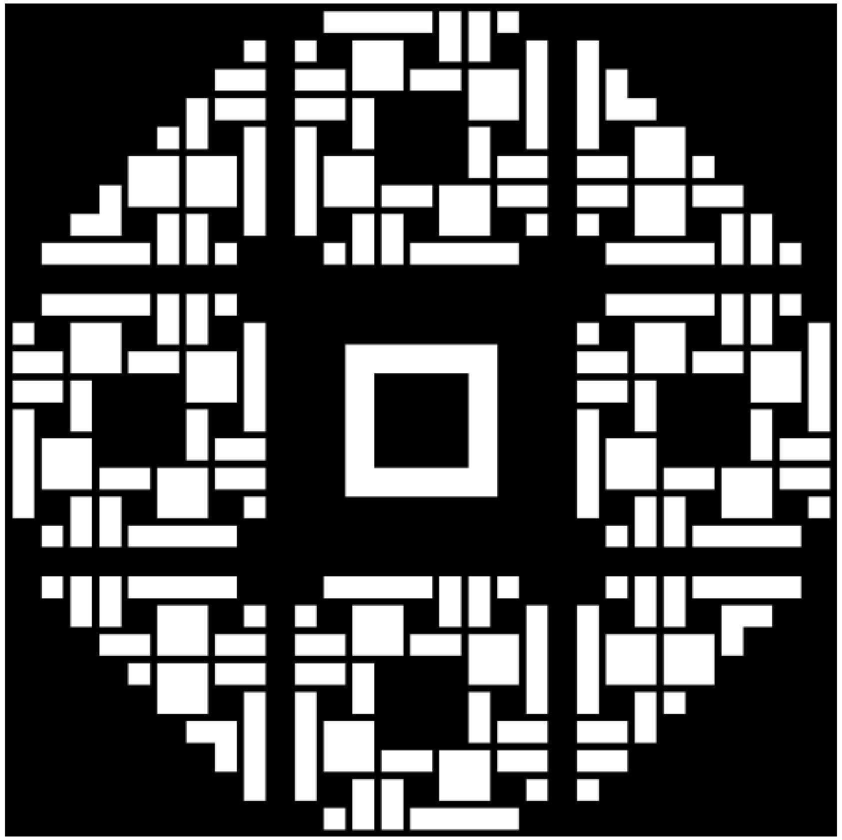


Figure 12.1 Design for a village in Algeria by Ricardo Bofill, with space black and buildings white.

in Figure 12.2 are based on some variant of the principles of enclosure, repetition and hierarchy. Similar principles can be detected in most of the schemes in the book. The 'enclosure-repetition-hierarchy' method became, at some stage of our recent past, it seems, a kind of international style of spatial design.

But the idea of the localised enclosure is not new. Nor is it dead. It was a common form in those precursors of modern housing, the 'philanthropic' estates of nineteenth-century London.[11] It was proposed by Le Corbusier as the fundamental, local unit of the new city in *La ville contemporaine*.[12] It was proposed by the late Greater London Council as the basis of 'good housing layout'.[13] Indeed, one of the curious things about the enclosure concept is the number of times it has been proposed as a new idea to remedy past errors. It might not be too fanciful to argue that it has acquired the kind of protected status we assign to concepts of unquestionable virtue, like 'motherhood' or 'community'. It encapsulates a moral imperative in a spatial idea.

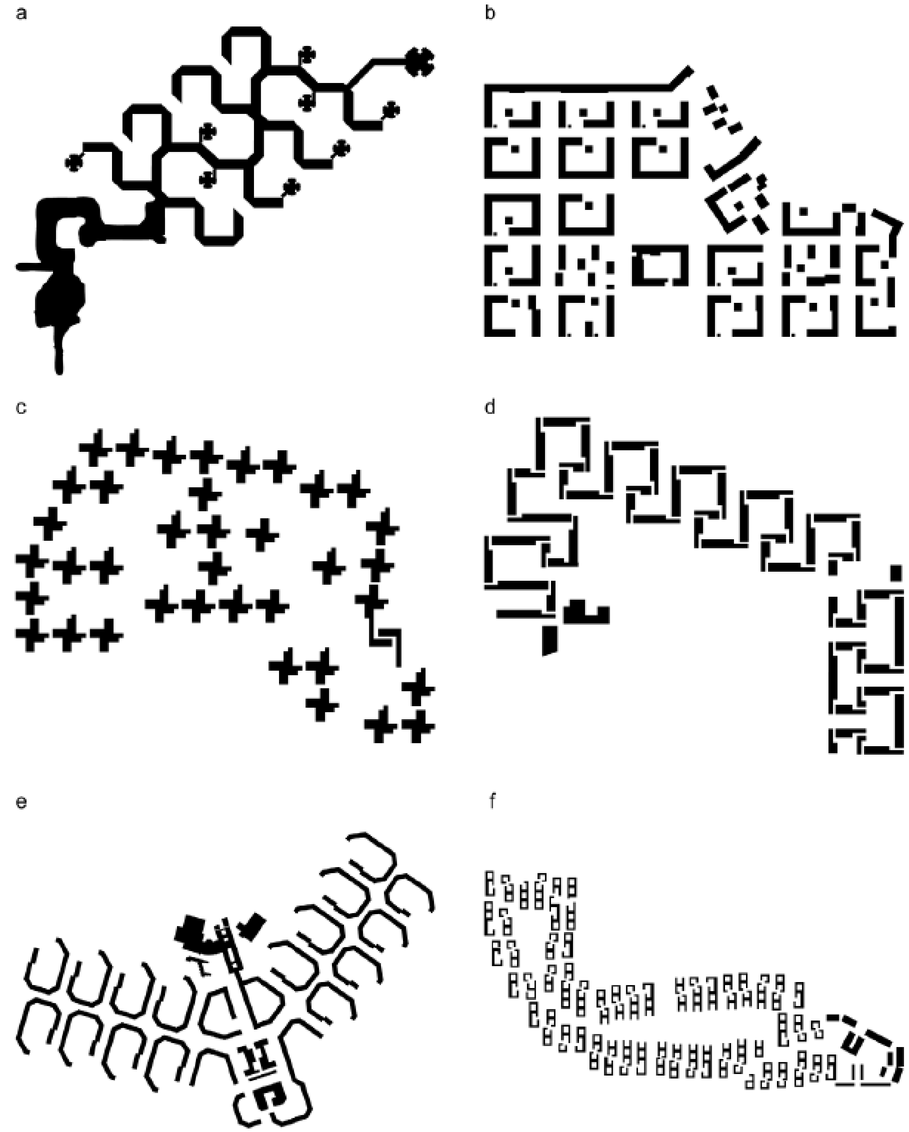


Figure 12.2 An international selection of housing schemes: (a) Dimona, Israel; (b) Bruket, Sandviken, Sweden; (c) Oriental Masonic Gardens, New Haven, USA; (d) Pollards Hill, England; (e) Steilshoop, Federal Republic of Germany; (f) Basildon, England (after Kirschenmann and Muschalek 1980).

My contention in this paper is, however, that the enclosure is not the answer to the urban problem, but the problem itself. Its indiscriminate use has been responsible for the creation of the fragmentary, unintelligible and largely underused spaces which form a significant proportion of our urban environment today. My aim in this paper is to make this argument precise by using a method of urban analysis – space syntax[14] – to show first how intelligibility and continuous use was created by urban space in the past, then show how the unintelligibility and underuse of much modern space arises from the uncritical use of over-localised concepts like 'enclosure'.

I will also show that there is now preliminary evidence that such environments may increase vulnerability to certain types of crime, even though we commonly think of them as 'defensible'. In conclusion, I will argue that to rehumanise our urban environment we must restrict the use of the enclosure concept to those places where it is genuinely applicable, and for run-of-the-mill public space reestablish the idea of open, outward-facing layouts, with intelligibility and integration given priority over exclusion and group territory.

Urban space isn't about enclosure

The enclosure concept claims legitimacy above all from the urban past. Now it is perfectly clear that there are enclosed spaces in our historic towns, and that the sense of enclosure can indeed be a pleasurable thing. But the belief that it is the basis of urban form is quite wrong.

Figure 12.3, the French town of Apt, is a fairly typical case of a traditional urban pattern. First, very little of the open space can be described as enclosed in the localised sense of Figure 12.2. Admittedly, all parts of open space are shaped and defined by their relation to building entrances; but, equally, all parts are related by lines of sight and access to the larger-scale space structure. Even the 'squares' which are the most obviously 'enclosed' spaces (and for the most part they are nineteenth-century interventions) have the important additional property of being also strategic spaces from which a good deal of the larger-scale space structure of the town can be seen.

Figure 12.4 shows this graphically. For each space, I have drawn a shape comprising all the space of the square, plus all the space that can be seen, and gone directly to, from any point in the square. The shapes almost join up to form a continuous structure, and with the exception of space D (a comparatively recent space which failed to develop as a market square), there is a clear relation between the size of the square and the size of the overall shape visible from the square. We might say that squares which are bigger 'locally' are also bigger 'globally'. Or that a space which is well defined as a local enclosure is also well defined in terms of the global structure of the town. This is quite unlike the localised concept of enclosure to be found in Figure 12.2.

Figure 12.3 The French town of Apt, Vaucluse, with buildings shown in white and open space in black.

The town also differs from the modern environments of Figure 12.2 in that there is virtually no morphological repetition anywhere. On the contrary, each part seems unlike all the others, to the point where one wonders how such a collection of idiosyncratic spaces can form a coherent system. As with 'repetition', so with 'hierarchy'. Only by stretching the meaning of the word to its limits can we detect anything resembling the modern idea of spatial hierarchy in the town.

But there is an even more fundamental difference between the old and new plan types. The new plans of Figures 12.1 and 12.2 are intelligible from the air, as plans. But if we try to move around them, we quickly lose all sense of where we are. The similarity of the parts, and their predominantly localised reference points, guarantee that, on the ground, they lack intelligibility. The old town plan has the opposite properties. From the air it appears disordered, in that it lacks the kind of regularity we have come to identify as urban order. But on the ground, it has a degree of natural intelligibility which means that we do not need signs to tell us where to go, or warnings to tell us when we are straying from the beaten track. We know how to read the town, and we know how to use it.

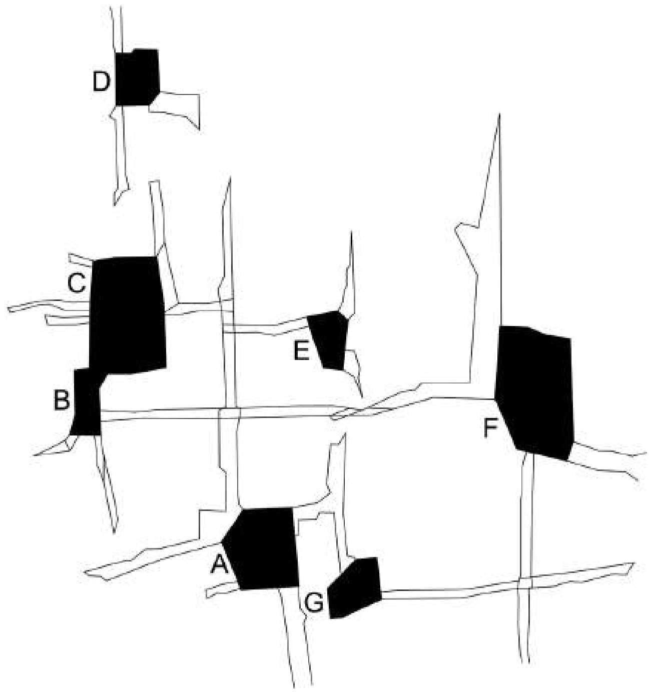


Figure 12.4 The 'convex isovists' of the squares of Apt, showing how 'local' enclosed spaces are related to the 'global' structure of space in the town.

Towns as deformed grids

How can this be? The answer, of course, is that traditional towns had forms of order built into them which gave them this intelligibility and workability, even though the concepts of urban order we have are quite incapable of expressing this order. This order is not mysterious. It is as well defined as any other kind of order. But it is different, and requires us to stop imposing our preconceived ideas on space and try to develop forms of analysis which are sensitive to these subtler types of order.

We can begin by noting that the great majority of towns built in human history are what I call '*deformed grids*'. By this I mean that the town has the general topology of a grid, being made up of a series of islands of outward-facing buildings, each surrounded by a ring of open space which forms part of an interconnected net. But its plan is 'deformed' in two senses. First, the length of spaces you can see changes as you move about; and second the width of spaces changes. We might say that the grid is deformed one-dimensionally, in terms of its lines of sight and access; and two-dimensionally, in terms of the changing 'fatness' of its spaces.

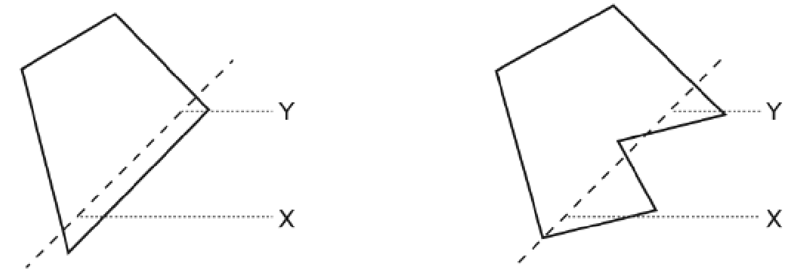


Figure 12.5 (a) A convex space, meaning that all points within the space can be joined to all others without passing outside the boundary of the space. In (b) the line from X to Y passes outside the space, which is therefore not convex.

These two types of deformation of the grid can be formalised and quantified. First, the two-dimensional variation can be analysed using the concept of *convexity*. As shown in Figure 12.5, a convex space is one in which all points are directly visible and accessible from all other points – there are no hiding places. Now, although it is quite easy to devise layouts which cannot be so analysed (for example, perfectly regular grids) it is, in practice, quite easy to divide most deformed grids up into their fattest and fewest convex spaces by simply requiring every point to be part of the fattest space it can. Since the boundary of each convex space will involve a vertex, the convex spaces can also be identified by drawing a line from each vertex to the nearest point on another island within the area defined by projecting the lines constituting the original vertex.

As soon as this decomposition is done, two key points immediately become clear. First, almost every convex space so defined, however small or narrow, has building entrances opening directly onto it. This is a social property as much as a spatial property. It means that every space from which all points can see all other points is also a space in which entrances participate in this 'all-play-all'. It also means that wherever you are in urban space, however localised, you are always under the potential surveillance of entrances. This is the opposite of modern enclosure-ism where, typically, entrances are deliberately concentrated onto a few spaces, to give these spaces identification and surveillance, leaving all other spaces without entrances and defined only by blank walls. The urban principle of the *continuous definition* of open space by building entrances is systematically violated in modern environments.

But convex spaces are not only related to entrances. They are also related to each other by lines of sight and access which pass through several, and perhaps many, convex spaces depending on which part of the settlement we are in. Again, this principle is commonly violated in modern imitations of the urban past. Lines of sight tend to be reduced towards the single convex space. The implications of this are also clear. It is the linking of convex spaces, through lines of sight, that give the deformed grid its sense of operating at two scales at once. Wherever you are, you are aware of the local space you are in through the convex organisation. But at the same time, you are aware of the global system of space by virtue of the lines of sight and access which connect directly to it. This way of operating at two scales at once is, I suggest, the single most important property of the deformed grid type of urban space structure.

Integrating cores

Just how important the linear or *axial* structure of urban space is can be seen by using it to develop a mathematical analysis of the global structure of the town. First, we construct an *axial map* by drawing the fewest and longest straight lines which pass through all the convex spaces of the settlement. Figure 12.6 is such a map for Apt. We can then see that to pass from each line to any other line, we must pass through a minimum number of intervening lines, or segments of lines. If we then count up the minimum number of lines we must use to go from each line to every other line, it turns out that all the sums will be different. This gives us the basis for a measure of *integration*: the fewer intervening lines which need to be passed through to go from a line to every other line, then the more integrating that line; the more, the more segregating. The computer can, in this way, be used to calculate an *integration value* for each line.[15]

We then return to the axial map and draw in the most integrating lines in order. If we stop at, for example, 10 per cent, then the resulting diagram we call the *10 per cent integrating core* of the settlement. Figure 12.7 is the 10 per cent integrating core of Apt, shown in heavy black, with the 50 per cent most segregated lines shown dotted. The core takes a form typical of many types of town or urban area, which we call a *deformed wheel*. A small semi-grid of lines in the heart of the settlement (the hub) is linked in several directions (the spokes) to lines on the periphery of the settlement (the rim), which also form part of the core. The segregated areas are found in the interstices formed by this wheel.

The integration core is probably the most important *deep structure* of the town plan. Its structure will vary from one type of town to another, but can usually be described as some part of the deformed wheel core (which we therefore believe to be fundamental). For example, we find *covering* cores (hub and spokes without the rim), *centralised* cores (hub only), *peripheral* cores (rim only), *penetrating* cores (one spoke and part of the rim), *linear* cores (one series of lines) and so on. Cores of any of these morphological types can then be localised in one part of the plan or globalised in the plan as a whole; *shallow* or *deep* in the plan; *fragmented* or *unified*, and so on.

Why should towns and urban areas have such different kinds of 'integrating cores'? Because, as the old adage says, towns are mechanisms for generating contact. But they do not all do it in the same way. The adage needs to be qualified:

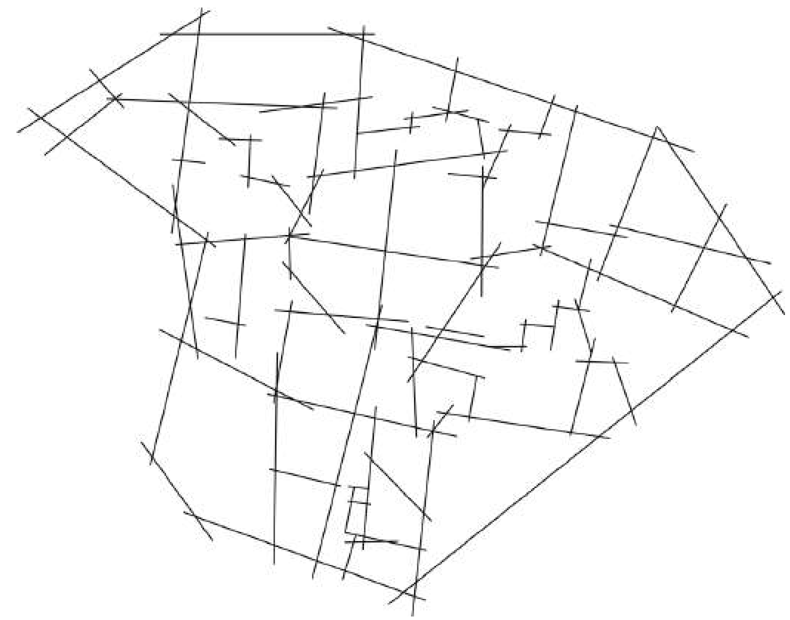


Figure 12.6 Axial map (or 'line diagram') of the town of Apt, made up of the fewest and longest straight lines that can be drawn through all the convex spaces of the plan.

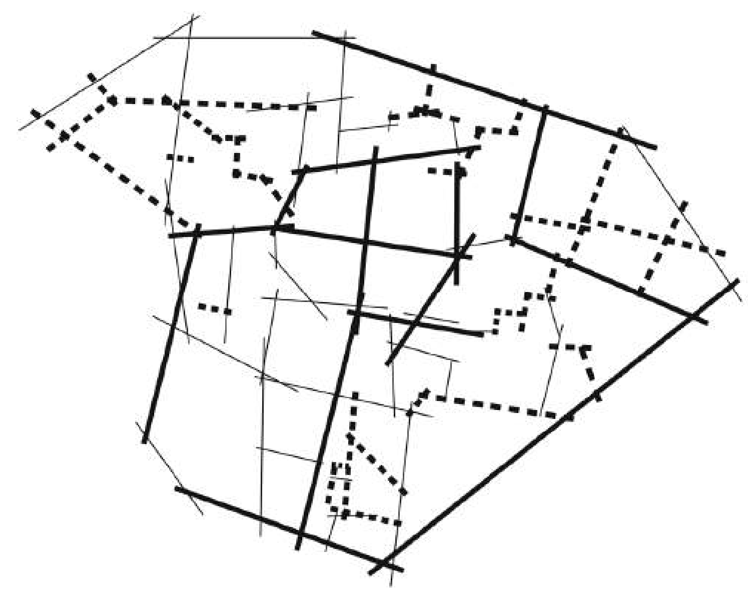


Figure 12.7 Axial map of Apt with the 10 per cent integration core shown in heavy black and the 50 per cent most segregated lines shown dotted.

how towns generate contact, and *what* kind of pattern of contact they engender is critical, and this is largely created by the way in which the town plan creates a pattern of integration and particular shape of core.

For example, the deformed wheel core type seems to exist to access strangers through into the heart of the town, while ensuring that the natural movement of inhabitants to, from and between the more segregated zones within the towns continually intersects the spaces used by strangers. This creates a strong natural 'probabilistic interface' between inhabitants and strangers in the town. In contrast, towns like Ghardaia in the Algerian Mzab restrict the global core and the movement of strangers to well-defined peripheral areas, and segregate large areas of the town for the more exclusive use of inhabitants.[16] Others again, like Venice, create a much more localised structure of dispersed cores, with an overall core concentrated in central areas.[17]

Encounter fields

These arguments link cultural variation in urban spatial forms to a common underlying principle: that the pattern of movement in a town is a function of its pattern of integration. This is, in fact, only conditionally true. It depends on other factors, the most important of which is another 'syntactic' property of the plan which we call 'intelligibility', discussed below. Even so, the relation between the pattern of integration and the pattern of movement is fundamental to understanding how towns and urban areas work as 'mechanisms for generating contact'; and the fact that 'integration' is a quantitative concept, derived only from a mathematical analysis of the plan, means that this is a testable proposition.

This can be clarified through the concept of the *encounter field*. The encounter field is the natural pattern of background space use and movement created by the town plan and the distribution of buildings within it. Research has led us to the firm conclusion that it is the structure of the town plan itself that is responsible for the bulk of this natural 'background' pattern of space use and movement, and that this decides the general level of use of the constituent spaces of the plan. The location of facilities and 'magnets' is important, but it is less influential in creating overall levels of space use than the town plan itself.

This proposition that the pattern of movement is fundamentally a function of the plan itself has been extensively tested by observation. The technique is simple. By systematically observing the pattern of movement in a set of spaces, we can assign to each space an *encounter rate*. We may then statistically correlate encounter rates with integration values (and with any other spatial parameters) and, obviously, the stronger the correlation, the better the prediction of the movement pattern from the spatial structure.

Figure 12.8, for example, is a map of the inner London area of Barnsbury, and Figure 12.9 its axial map with integration and segregation marked in the usual

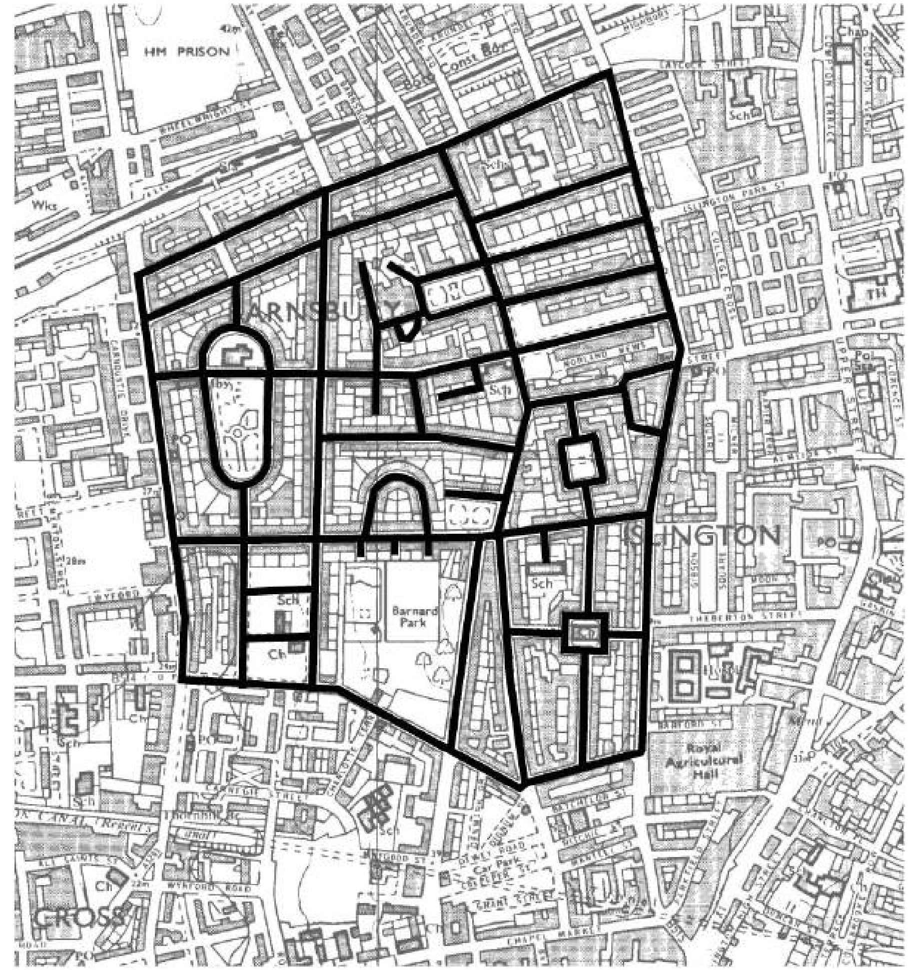


Figure 12.8 The area of Barnsbury, North London, with open spaces shown in black.

way, showing a 'covering core' centred on the 'village' (the short line marked 1) and the residential squares as the more segregated spaces. Figure 12.10 is then a scattergram setting encounter rates of lines along the horizontal axis and integration values along the vertical. The correlation of 0.8 (on a scale from 0 = no relation to 1 = a perfect relation) shows how consistently the encounter rate varies with the degree of integration of the line.

Research has shown that this type of result is normal in street-based urban areas: integration values are strong (and the best) predictors of the encounter rates of individual spaces. Because integration is a measure which relates each space to every other space in the plan, this implies that the encounter rates of individual spaces are, in the main, a function of their position in the 'global' structure of

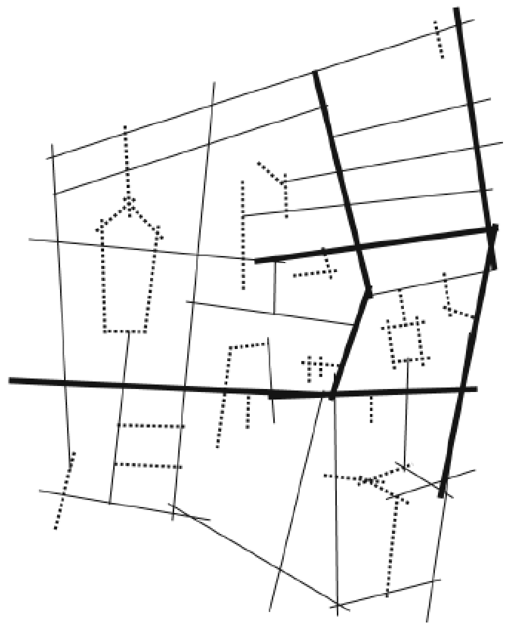


Figure 12.9 Axial map of Barnsbury, showing its 10 per cent integration core in heavy black and 50 per cent most segregated spaces dotted.

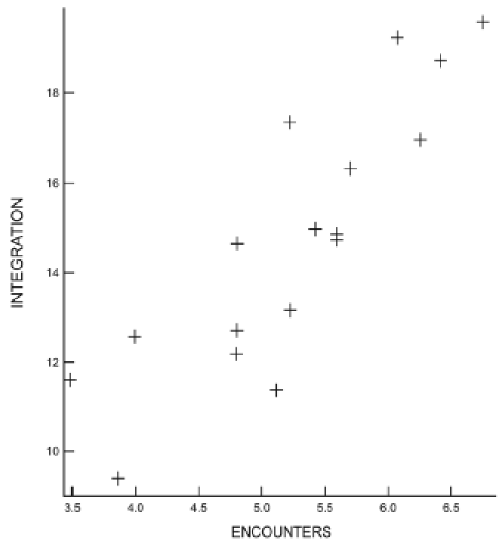


Figure 12.10 Scattergram of numbers of moving people (horizontal axis) plotted against integration (vertical axis), showing a correlation of r = 0.8004. Note that in this case the reciprocals of integration values are used so that higher values mean more integration.

the plan, not the more 'local' properties of the space. In other words, how the urban 'part' works depends on its relation to the urban 'whole'. This can be found even at a much larger scale of analysis. Figure 12.11, for example, is an axial map of the street pattern of a much larger urban area, of which Barnsbury is a part, Figure 12.12, the same with the addition of the ground level of housing estates built over the past three decades, and Figure 12.13 a scattergram plotting integration values against encounter rates for a much larger sample of spaces taken from six separate studies distributed throughout the area. The correlation is still remarkably good, although the scatter shows more 'spread' than in the more localised area.

However, this overall result conceals more important effects that can be found by looking at the six studies independently. First, we may look carefully at the differences between Figures 12.11 and 12.12. Those estates added in Figure 12.12 vary enormously in their layout, but all involve the enclosure theme in some way. One is an early, post-war, Modernist estate with a large-scale court form, another a 1960s slabs estate, again arranged loosely in separate courts, another a low-rise 1970s housing estate in small courts, another a low-rise development around a 'village green', another a trend-setting, neo-vernacular, pseudo-village full of 'focal' spaces, and so on.

In spite of the differences in style and geometry, the axial map shows that all the estates have certain common properties, which make them obtrusive in

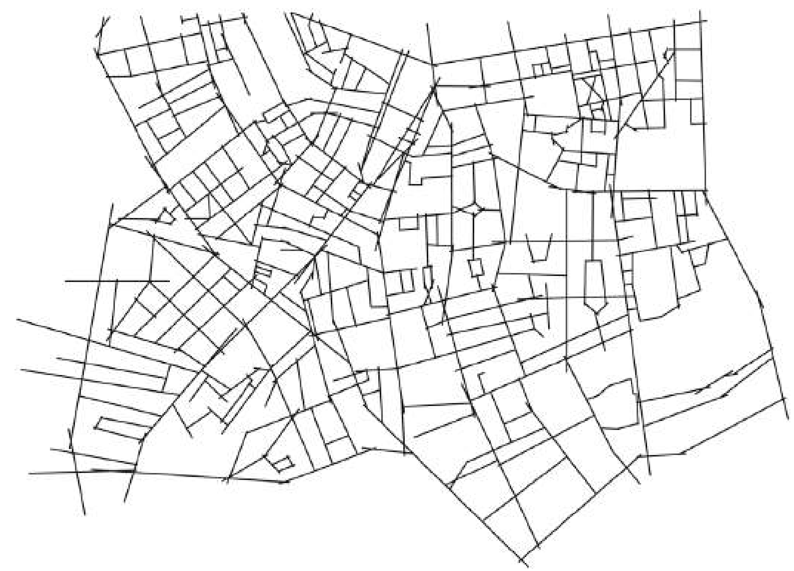


Figure 12.11 Axial map of street pattern of the large study area in Islington.

Figure 12.12 Figure 12.11 with the addition of estates built in the last three decades.

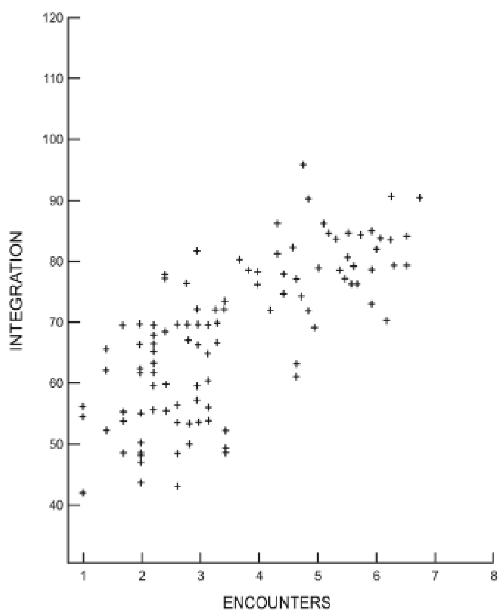


Figure 12.13 Scattergram of encounter rates against integration for all six study areas in Figure 12.12.

the axial map. First, and most obvious, is a radical *reduction* in spatial scale, even where the buildings are quite massive. This is an almost universal property of modern housing. Second, the estates seem to lack any sense of an internal spatial structure, either in plan or on the ground. This turns out to be the case. Syntactic analysis shows the estates have, almost uniformly, largely peripheral integration cores. They integrate from the outside. This is reflected in the pattern of movement, which concentrates on the exterior and falls off rapidly as you move deeper into the estate. The encounter field makes sense only in terms of people moving in and out of the estate by the shortest possible route.

But the estates also share a third, less obvious property. The spatial patterns appear to lack intelligibility *either* from the air as plans *or* on the ground as places to move around. Our research has suggested there may be a powerful way of measuring this. Each line has a certain number of other lines that intersect with it. Call this number the *connectivity* of the line. Connectivity is a property of a line which can be *seen* from the line. In this sense it is a *local* property. The integration value of a line, on the other hand, cannot be seen from the line, because it is a global property. It requires knowledge of the system as a whole, most of which cannot be seen. If we then correlate the connectivity values of lines with their integration values, then we will have an index of the degree to which the spatial information available visibly from a line is a reliable guide to the importance of that line in the system as a whole.

This correlation, expressed as a value between 0 and 1, we call the 'intelligibility' of the system. Analysis of a large number of examples from different parts of the world has shown that intelligibility, defined in this way, is a key property of the spatial structure of towns. It has also been shown to have a powerful effect on the encounter field. Roughly, the more intelligible the urban area, the better will be the prediction of encounter rates from integration values. In traditional urban areas both tend to be high. This confirms intuition. We feel they are intelligible. And we feel that there is a natural relationship between the presence of people and the spatial pattern.

Lost properties

In the estates we have studied, both of these relations commonly break down. The internal intelligibility of the spatial layout falls, and the degree to which the pattern of integration predicts encounter rates falls with it, both to much lower levels than in the surrounding urban areas. The encounter field within the estate does not relate to the internal structure of space. It only makes spatial sense in terms of movement into and out of the estate by the shortest available routes. The result is the loss of both of the key elements of the 'urban' relation between space and people: the sense that the spatial layout, as a whole, is intelligible from its parts; and the sense that the spatial layout creates a predictable pattern of encounter.

But there is a third, even worse outcome. In almost all estates, whatever the housing density, the overall encounter rate drops from an average of around 2.6 people per hundred metres/minute of walking time in street-based, residential street areas (not counting shopping streets) to somewhere between 0.4 and 0.7 inside estates. The effect is that, whereas in streets you are in contact with other people most of the time, on estates you are on your own in space most of the time. Put another way, the daytime encounter field in the estates turns out to be like night-time in ordinary urban streets. In terms of their naturally available encounter field, people on these estates live in a kind of perpetual night.

Other, less obvious, properties of the encounter field are also lost. For example, in street-based areas, the pattern of space use by adults and children, or women and men, is more or less similar, giving a strong degree of natural interfacing between these different categories. This can, again, be expressed as a correlation coefficient. A high correlation between the encounter rates of adults and children for spaces will mean that adults and children are in a constant natural interface with each other. In many estates, this correlation becomes very weak, and in some it even goes negative, implying an effective separation of the encounter fields of adults and children.

Vulnerability to crime

What the long-term, social effects of these 'lost properties' may be can only be conjectured at present. But at least one important social variable can be brought into focus by these methods: vulnerability to crime. Work on the relation between spatial layout and crime location is now being carried out by Lena Tsoskounoglou, a doctoral student in the Unit for Architectural Studies, University College London.[18] The ability to express configurational properties of layouts quantitatively has allowed her to approach the relation of architecture and crime vulnerability in a new way. Each location, in an estate or urban area, can be assigned various spatial parameters expressing different aspects of its relation to the layout as a whole. Different types of crime can then be plotted on the layout, and thus be assigned spatial values. From there, it is a matter of exploring how far different types of crime tend to be associated with different spatial values.

For example, Figure 12.14 is the axial map of Barnsbury, with integration and segregation marked in the usual way. All burglaries reported to the police during a recent 12-month period are marked by heavy dots, giving an overall rate of 5.4 per cent of buildings burgled in one year. Intuitively, it looks as though more segregated locations have higher rates. Can this be tested?

There are two ways of doing this. First, each dwelling is assigned the integration value of the line (or lines, if there is also a back entrance) onto which it opens, and the mean integration value of burgled dwellings and unburgled dwellings are calculated separately. The result, in this case, is that burgled dwellings

Figure 12.14 Axial map of Barnsbury with burglaries reported in a 12-month period shown by dots.

have a mean integration value of 0.6936 whereas unburgled dwellings have a mean of 0.6539. Since low values mean more integration (that is, less depth), this result shows that segregated dwellings are more likely to be burgled than integrated dwellings. A significance test on this difference shows that it is statistically 'highly significant' (less than a 0.05 probability of having occurred by chance).

Second, we can calculate the *rate* of burglary for a line as the number of burgled dwellings over the total number of dwellings on that line. We can then correlate these rates with the integration values of the lines. If we do this for Barnsbury, then at first we find only a weak relationship between the two, showing that vulnerability to burglary increases slightly with increasing segregation. However, an inspection of the scattergram shows that this apparently weak result is entirely due to the fact that over half the total spaces had no burglary during the 12-month period, and these are spread throughout the range of integrated and segregated spaces. A good deal of this must be due to chance. In any 12-month period, there will only be enough burglary to give positive rates for about half the lines. If we then look only at the lines on which burglary occurred in the 12-month period, and plot these against integration, then the result is the scattergram in Figure 12.15, showing a very strong, and statistically significant, relation between integration and burglary rates. The more integrated the space, the lower the burglary rate, the more segregated, the higher the burglary rate.

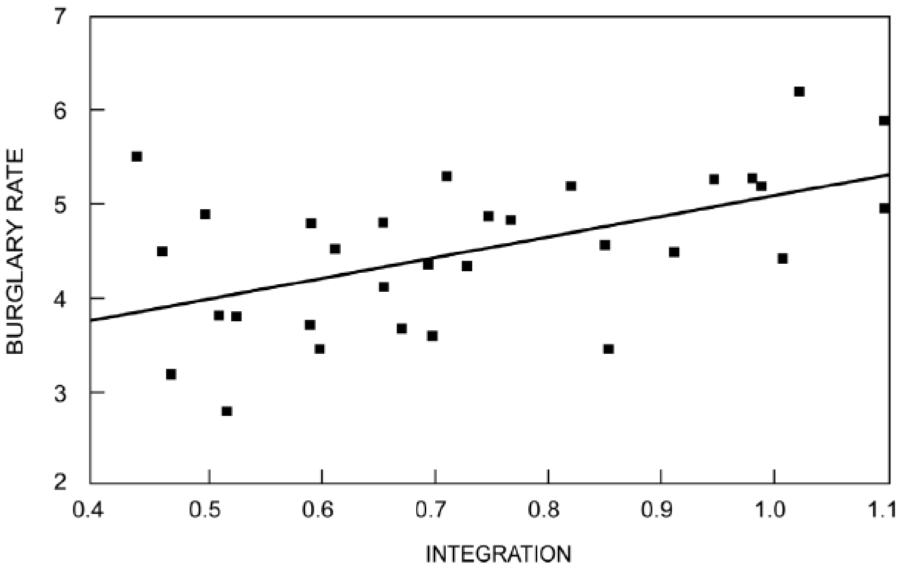


Figure 12.15 Scattergram of burglary rates (vertical axis) against integration (horizontal axis) for Barnsbury during one 12-month period. Note that in this case ordinary integration values are used, meaning that a low value means strong integration.

This result must be treated with great caution, of course, because of the fact that a single 12-month period does not yield enough data to test this for all lines in the system. Data taken over several years would be required to eliminate this problem, and work is now in progress to try to build this up. The result is, however, very suggestive, and taken in conjunction with the first result, strongly suggests that burglary risk *rises* with the degree of segregation, that is with the spatial variable that *reduces* the encounter rate of the space, and falls with integration, which produces higher encounter rates.

In other words, Jane Jacobs's instinct seems to have been right when, in *The death and life of the great American cities*, she associated higher encounter rates with lower crime risk.[19] Does the same apply on estates, or is it better there to segregate and reduce the encounter rate, as Oscar Newman and Alice Coleman suggest?[20] Figures 12.16 and 12.17 are the plan and axial map of the ground level of part of one of the estates in Figure 12.11, the Marquess Road estate. The overall rate of burglary at 8.9 per cent is much higher than for the more affluent Barnsbury area but, also, there are interesting differences in the levels of the estate. The rate for the upper walkway level (not shown in the figures, but it is less labyrinthine than the lower level) is only half the mean rate for the whole, while those parts of the ground level which are slightly raised by ramps or steps (but connect to the ground level rather than to the upper level) have double the mean rate for the whole.

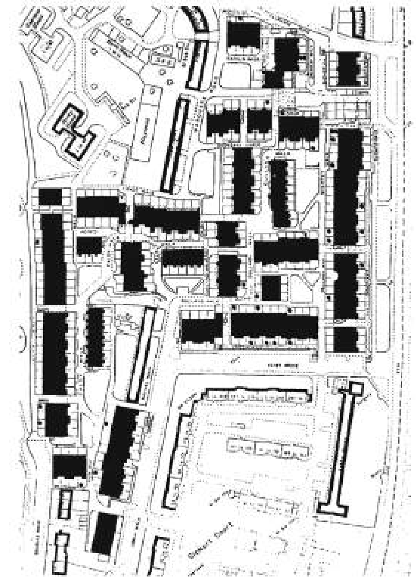


Figure 12.16 Plan of the ground level of the Marquess Road estate in Islington.

Figure 12.17 Axial map of the ground and ramped levels of the Marquess Road estate with burglaries reported in a 12-month period shown by dots.

For the ground (and ground connected) level the mean integration value of burgled dwellings is 0.8746 while for unburgled it is 0.8078 – again burgled dwellings are more segregated than non-burgled and, once again, this difference is statistically highly significant. The scattergram for burglary rates against integration values for burgled lines, again, shows that rates increase markedly with segregation, with a correlation of 0.59. At the upper level, where the rate is much lower, the pattern is less clear. The mean integration of burgled dwellings is, in fact, slightly lower than for the unburgled, but the difference is not statistically significant (there are only 11 burglaries). We still find, however, that burglary rates rise with segregation.

A visual inspection of the axial map of the ground-related level suggests that another spatial factor is implicated in the distribution of burglary. Figure 12.18 is the axial map of the estate with line segments, onto which no dwelling opens, drawn in heavy black. It is immediately clear that the strongest sequences of such 'blank wall' lines occur at all the entrances to the estate. A more careful look will show that nearly all burglaries occur on the first line in from the outside on which entrances occur. Burglars, it seems, like to enter and leave by routes which do not take them past entrances.

Two figures powerfully confirm this pattern. If we calculate the mean burglary rate for all such 'first' lines with entrances from the outside, then we find a rate of 21 per cent. If, on the other hand, we look at the rates for the lines that are completely protected from the outside by lines with entrances, then the rate is around 2 per cent. No such pattern can be found if one tries to look at the surveillance of individual lines. It is not, it seems, the surveillance of the space onto which you open that keeps you safe, but the potential surveillance of the routes to your space from the outside.

The consideration also seems to play a role in the very different spatial pattern of the second estate, the Andover estate, the ground level of which is shown in Figures 12.19 (plan) and 12.20 (axial map with burglaries). Again, the estate is on two levels, and once again the upper, deck access level (which is not really an upper-level walkway system since it continually requires one to return to the ground level to move about) is rather safer than the ground floor (6.25 per cent as against 7.87 per cent). The estate is unusual in that it has an integration core which passes through the heart of the estate, linking it to the outside in several directions. However, this integration core is almost entirely made up of entrance-free lines, so that there is no direct interface between dwellings and people moving through the estate.

Again, we find a different pattern for the ground and upper levels. On the ground, the mean rate for burgled dwellings, at 0.5936, is marginally more integrated than for unburgled dwellings, at 0.6147, although, again, this is not statistically significant. This could be because, as we have said, the integration core passes through the heart of the estate almost entirely unrelated to dwelling entrances. At the upper level, however, there is a strongly significant difference

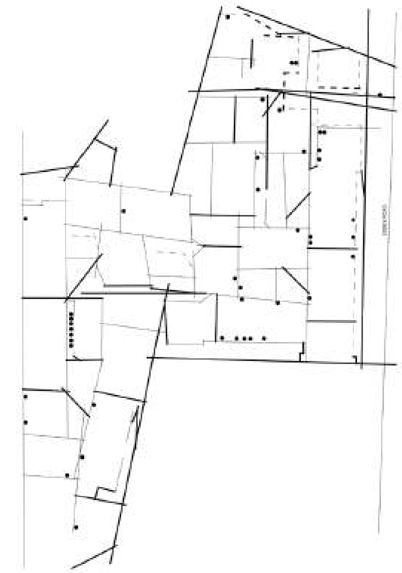


Figure 12.18 Axial map in Figure 12.17 showing spaces without building entrances opening onto them in heavy black.

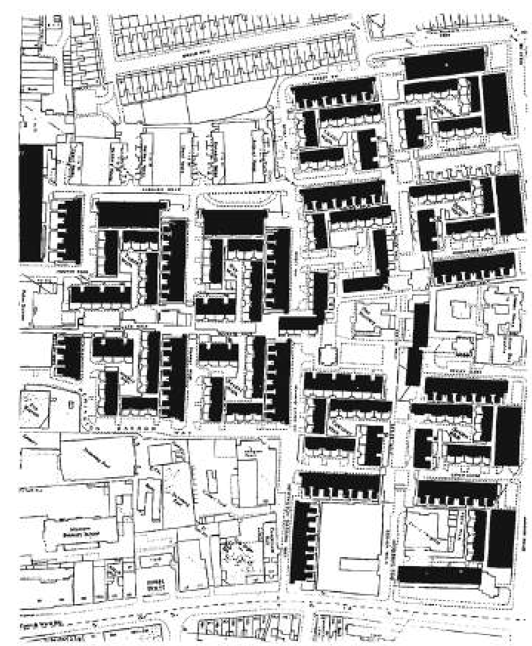


Figure 12.19 Ground-level plan of the Andover estate, Islington.

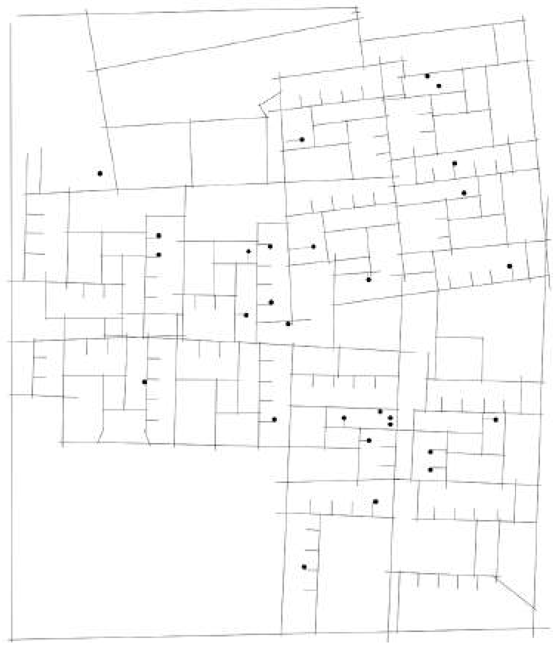


Figure 12.20 Axial map of the ground level of the Andover estate with burglaries during a 12-month period shown as dots.

the other way: burgled dwellings, at 0.9807, are significantly more segregated than unburgled dwellings at 0.9361. The scattergram for burglary rate against integration for the burgled lines on both levels (the upper level cannot be dealt with as a separate system in this case) again shows a strong tendency for rates to increase with segregation, with a correlation of 0.55.

Taking the three areas together, then, for a total of 2,816 dwellings, the 211 burgled dwellings have a mean integration of 0.780 while the unburgled are 0.759, that is, burgled dwellings are markedly more segregated. This difference is clearly substantial, although, for technical reasons (the three systems have not yet been assembled into a single spatial database, because of the size this system would need to be) it has not yet been possible to compute a significance test for it. It leaves little doubt, however, that in general, the properties of urban space that emerged as the key ones in our studies – those of integration and those of continuous relations to building entrances – are both critically bound up with the pattern of burglary.

The results also cast quite fundamental doubts on the whole concept of 'defensible space', at least insofar as one of the main assumptions behind it is that the elimination of natural movement and encounter within housing estates will increase safety. Advocates of defensible space, from Newman onwards, seem to believe that the criminals seeking victims are part of the passing crowd and that strangers are therefore, in principle, dangerous. Something like the opposite appears to be the case. The natural presence of people may be the primary means by which space is policed naturally. The more you eliminate this, then the more you create danger once a potential criminal has appeared on the scene. It is true that people behave more 'territorially' in segregated spaces. The more segregated, the more likely one is to question the presence of strangers. But this is associated with feeling more unsafe. No one feels the need to question strangers passing down a street. On the contrary, their natural presence increases the sense of security. It may also, it seems, increase actual security – although much more research is needed before these suggestive findings can be turned into unequivocal results.

So how should we rehumanise housing?

The design strategies that are implied by these new research findings can be set out at two levels: at the level of the 'rules of thumb', which ought to be adequate in the case of smaller-scale design problems; and the level of the proper 'preparatory study' before design is initiated, which becomes more necessary as the size and complexity of the design task increases. In either case, what is proposed below is intended to deal as much with the problem of modifying existing housing areas and estates – and avoiding expenditure on modifications which will not improve the situation – as with the design of new developments. The rules proposed would apply as much to shopping areas and mixed areas as to housing areas.

First, the rules of thumb. We use the term 'line diagram' for the axial structure of a proposed scheme, and 'space diagram' for the convex structure. The following can be proposed:

- If it is intended that the new, or modified, area should relate effectively to the surrounding area, make sure that the 'line diagram' of the sketch design links the heart of the scheme with the surrounding area in several directions with lines that permit both visibility and direct access. These key lines should not pass right through the scheme, but reach important destinations within the scheme with one line, before changing direction to another. These lines should take into account the dominant lines that already exist in the surrounding area, not by continuing them, but by redirecting them. In effect, this first stage of design is the design of an integrating core of the type desired.
- In developing the line diagram of the scheme, make sure that all lines are, at most, two lines deep both from the outside and from the integrating core, with no more than an occasional line three steps deep.
- Make sure that any rings in the line diagram (that is, choices of routes), including those that form rings with the surrounding area, are related to the integration core. If rings are too segregated, then it creates access without adequate use. Choices of route are good, provided they are all adequately used.

- Make sure that all spaces in the space diagram, however small or narrow, have building entrances opening directly onto them. If this is impracticable for small spaces, then avoid creating those small spaces. Avoid clustering too many entrances onto too few spaces, and concentrate instead on trying to ensure that every part of the scheme is in touch with entrances. Particularly sensitive, in this respect, are the spaces leading into the scheme from outside. These should always be related directly to building entrances.
- Make sure that spaces in the space diagram have links of visibility and direct access through the line diagram to the larger-scale structure of the scheme. The 'isovist' (what can be seen and gone to directly from a space) of a space should be roughly proportional to its size.
- Make sure that the orientation of building facades, and their entrances, are such as to clarify the line and space structure of the scheme. For example, lines of sight striking buildings at open angles will suggest further movement possibilities; and marking important moments in the spatial structure with key facades will aid intelligibility and memorability.
- Avoid over-enclosing spaces, except where this deliberately reflects the place of that space in the overall spatial syntax of the scheme.
- Avoid repetition and simple geometrical permutations as far as possible; local differences aid global intelligibility if they are handled well.
- Avoid the over-hierarchisation of space; a range of rather more integrating, and rather more segregating, spaces is enough to differentiate the parts of the system into busier and quieter zones, and will avoid creating space that is empty for most of the time.

If the scheme is large or complex enough to merit a proper preparatory study, then:

- The area surrounding the proposed scheme, or modification, should be analysed syntactically to establish its existing structure. This should be done both with, and without, whatever is on the redevelopment site at present. An area with a diameter of about 1 kilometre is usually adequate to give a firm basis to the study.
- The existing pattern of space use and movement in the area should be studied through a sample of spaces selected to cover the range of space types in the area: integrating and segregating, shopping and residential, and so on. The various types of encounter rates should be correlated with the spatial analyses to give a clear picture of the functioning of the area, as it stands, and how it relates to the spatial structure.
- Use the analysis of the area to generate alternative sketch layouts, then insert these into the computer model of the area and re-run the analyses. This will show the effect of the area on each scheme and the effect of each scheme on the area. It is also possible, at this stage, to simulate likely movement patterns in each of the design alternatives, using the knowledge gained in the study of the existing area.
- Use this analysis to locate key buildings and facilities which need a particular type of relation to the surrounding area, and to the internal structure of space on this site.
- Then proceed, as in the 'rules of thumb' section, but at each stage check the design against the computer analysis by simply adapting the computer model (a few moments' work). In this way, intuitive design and systematic evaluation can proceed cyclically to generate a scheme that satisfies the given objectives.

Notes

1. J. Jacobs, *The death and life of great American cities* (Harmondsworth, Middlesex: Penguin, 1961).
2. O. Newman, *Defensible space: People and design in the violent city* (London: Architectural Press, 1972).
3. B. Hillier and J. Hanson, *The social logic of space* (Cambridge: Cambridge University Press, 1984), p. ix; B. Hillier, *Space is the machine: A configurational theory of architecture* (Cambridge: Cambridge University Press, 1996), p. 187.
4. L. Lees and E. Warwick, *Defensible space on the move: Mobilisation in English housing policy and practice* (Newark, NJ: John Wiley & Sons, 2022).
5. J. Hanson, 'Urban transformations: A history of design ideas', *Urban Design International* 5 (2000): 97–122, p. 116.
6. J. Peponis, C. Ross and M. Rashid, 'The structure of urban space, movement and co-presence: The case of Atlanta', *Geoforum* 28 (1997): 341–358, p. 341.
7. A. Legeby, 'Patterns of co-presence: spatial configuration and social segregation' (PhD thesis, KTH Royal Institute of Technology, 2013), p. 94.
8. L. Vaughan, C. E. Jones, S. Griffiths and M. Haklay, 'The spatial signature of suburban town centres', *Journal of Space Syntax* 1 (2010): 77–91, p. 77.
9. J. Hanson and B. Hillier, 'The architecture of community: Some new proposals on the social consequences of architectural and planning decisions', *Architecture et Comportement/ Architecture and Behaviour* 3 (1987): 251–273.
10. J. C. Kirschenmann and C. Muschalek, *Residential districts* (London: Harper Collins, 1980) (Originally in German as *Quartiere zum Wohnen*, Deutsche Verlags-Anstalt, 1977).
11. J. N. Tarn, *Five per cent philanthropy: An account of housing in urban areas between 1840 and 1914* (Cambridge: Cambridge University Press, 1973).
12. Le Corbusier, 'A Contemporary City (La Ville Contemporaine)', in *The City of Tomorrow and Its Planning*, translated and with an introduction by Frederick Etchells (London: John Rodker, 1929), pp. 217 and 221.
13. GLC, *An introduction to housing layout* (London: Architectural Press, 1978).
14. Hillier and Hanson, *The social logic of space*; B. Hillier, J. Hanson, J. Peponis, J. Hudson and R. Burdett, 'Space syntax: A different urban perspective', *The Architects' Journal* 178(48) (1983): 47–63.
15. For equations, see Hillier and Hanson, *The social logic of space*; Hillier et al., 'Space syntax'.
16. F. Salah-Salah, 'Cities in the Sahara: Spatial structure and generative processes' (PhD thesis, University of London, 1987).
17. C. Zeliotis, 'The spatial structure of Venice' (MSc diss., Unit for Architectural Studies, University College London, 1985).
18. The student's thesis was published 1995. E. Tsoskounoglou, 'Spatial vulnerability to crime in the design of housing estates' (PhD thesis, University of London, 1995) (eds).
19. Jacobs, *The death and life of great American cities*, for example, p. 43.
20. Newman, *Defensible space*, chapter 2; A. Coleman, *Utopia on trial: Vision and reality in planned housing* (London: Hilary Shipman, 1985), p. 147.

13 The architecture of the urban object (1989)

Commoditas, communitas and the virtual community. Introduction to 'The architecture of the urban object'

Laura Vaughan

'The architecture of the urban object', published in a special issue of *Ekistics* in 1989, is a central text in the space syntax canon, for its establishment of a social theory of how urbanity is both shaped by spatial configuration and is measurable as a mathematical property of the network. Opening his exposition by citing key texts from the disciplines of historical geography and urban history – notably situating his analysis in the anthropological research domain too – Hillier anchors his discussion in a deep understanding of the historicity of cities. This is crucial to his argument that takes issue with the sort of studies that read the surface of the built environment, ignoring how its materiality is not only a reflection of past cultures and embedded memories, but forms part of the everyday life of the city today. While in its closing sections the article dissects the reasons for the apparent failure of Modernist planning to create cities as social spaces, it does much more than that.

Arguing that architecture is comprised of construction, style and space (*commoditas*) leads Hillier to argue for a concept of a spatial culture: 'a distinctive way of ordering space to produce and reproduce . . . the principles for ordering social relations'. This is an important point in space syntax theory, refuting criticisms of its supposed architectural determinism by showing that analysis of linear representations of urban (and architectural) arrangements as configurations produces solid evidence for how spatial configuration can *give rise* to patterns of connections between people (not determine them). Here Hillier proposes the conception of spatial form as creating the 'field of probable – though not all possible – encounter and co-presence within which we live and move': through this defining an essential concept in space syntax theory, the *virtual community*. As Peponis and colleagues have explained, the term was coined in order to capture the 'social effect of space, quite independently of more conventional definitions of community as a function of consciously shared ways of life, or active interaction'.[1] This is a challenging concept, which Hillier has subsequently sought to pin down by arguing that it is measurable – by examining the relative density of movement flows, taking account of different categories of people.[2]

Indeed, the potentiality of urban life is itself an important point. First, the underpinning rules for how buildings, streets and street configurations take shape are, Hillier argues, governed by cultural rules regarding the interface not only between locals and strangers but also between different groups of people who live in the city: a set of community memberships that are not ordered but structured by the spatial configuration of the city. One of the most important contributions to urban theory is set out here: the argument that urban communities are not one-to-one reflections of a particular territory within the city but are comprised of myriad spatial (and in some instances transpatial) socio-spatial realisations in space. In a long passage that discusses Victor Turner's 'The ritual process', Hillier turns to the concept of *communitas*.[3] Life on the street, he argues, is analogous to the state of liminality that is created by religious ritual, which results in a temporary suspension of social differentiation. Thus, communities become undifferentiated in the public realm, forming 'the world of strangers, a form of community perpetually growing or shrinking' in response to the structure of space itself. This concept, in many ways a development of the configurationally significant *virtual community*, has become a touchstone in space syntax theory: the notion that co-presence is a vital aspect of urban life, and co-absence a sign of public space being pathological.

Hillier refers here to 'The architecture of community', in which he and Julienne Hanson argued that the role of the city is to organise difference, to 'reassemble what society divides'.[4] The city, they argued, creates a structured non-correspondence between community and space. In other words, the city street becomes a shared space for people who do not know each other, nor have any more in common than that they inhabit a common set of spaces. In the closing passages Hillier highlights how, in its emphasis on localism, modern city planning can lead to a critical loss of global order and, thus to an absence of the sense of urbanity. He shows that the relationship between the syntax of space itself and the virtual community is 'at the heart of the urban question'. It is this intrinsic urbanity of cities that is so important to the space syntax theory of urban space. Hillier illustrates how in traditional cities, different kinds of solidarity co-exist in space, finding different modes of localisation with the support of a type of configuration that supports structured integration. As subsequent articles will set out, the uniqueness of each place, the shared sense of place provided by each city's *genius loci*, becomes the next challenge for the theory to address: to capture how the relationship between pathway structure and human movement creates each city's own individual identity – or spatial culture.

The architecture of the urban object

Bill Hillier

Originally published: Hillier, B. 1989. 'The architecture of the urban object'. *Ekistics*, 56, 334–335. Special issue: *Space syntax: Social implications of urban layouts*, 5–21.

Introduction

'Architecture' is a strange word. We use it to refer to the appearances of things, as in 'baroque architecture' or 'seaside architecture', but we also use it to refer to the deep structure of things, as in 'the architecture of matter' or the 'architecture of the cell'. The expression 'the architecture of the city' thus has a certain ambiguity. It might mean the characteristic or appropriate appearance of buildings in cities. Or it might mean the deep structure of the city itself as a material object. This essay is about the second meaning: the 'architecture' of the urban object itself.

The city as architectural object

Two questions immediately arise: does such a thing exist? And if so, is its existence significant? Much recent theory on the origins and nature of urbanism suggests that even if it does exist, its significance is doubtful.[5] The city is, after all, so much more than a material artefact. It is economic processes, social relations, psychological states, cultural milieux and so on. Each offers a different way of studying the city and a different base from which to construct the theories we need in order to act upon it. It seems that the more we have become more aware that the study of the city involves the study of these dynamic processes, the more we have come to see the material form of the city as an epiphenomenon, a by-product of these dynamic processes. Its relevance has come to be seen only in terms of the clues that it can give to the form and nature of these processes, not as an essential constituent of them. In the extreme form of this argument, the material city vanishes, to be replaced by a despatialised network of communications and transactions.[6]

Architecture, defeated by the failure of Modernist social engineering, has found few effective ways to re-argue the cause of the material city. It has responded either with theories that reject the twentieth century and see the architecture of the historic city in terms of a lost symbolic tradition (for example Rykwert 1976), or vanished lifestyle (for example Krier 1979), or with theories that celebrate the architecture of appearances of the modern city but which deny any questions of deeper 'architecture'.[7] None challenge the power and dominance of the theories which de-materialise the global form of the city by linking the architecture of the city to its dynamic processes.

Yet all these views contain a paradox. If the city is distinctive in the richness and density of its social, economic and cultural milieux, then there must be some sense in which these arise from the material basis of the city. It is not enough to see the architecture of the city as a vague reflection of its social forms. If social forms are unique to the material city, then logic requires that there must also be some sense in which they are a product of the material city. Common sense affirms this. Cities of different cultural types and different scales embody different spatial identities. Our experience of these seems intrinsic to what cities are. We do not de-materialise our urban experience in the way in which we have de-materialised our urban theories.

What I sketch in this essay is a theory of the architecture of the urban object itself, that is a theory of the deep structure of the material form of the city, both as an autonomous reality in itself and as an essential constituent of the dynamic processes that make up the city. Aldo Rossi was, I believe, the first author in modern times to advance a theory arguing that the material form of the city is intrinsic to its sociological, cultural and psychological reality.[8] In this paper I take a position that also goes well beyond Rossi's theory by trying to demonstrate empirically as well as theoretically the autonomy of the urban artefact, and then go on to argue that it is only by understanding this autonomy that we can understand why the material form of the city is an intrinsic aspect of its social existence.

The laws of the urban object

I will argue that to understand the city we must *first* understand its material form, and most especially its spatial form; and that we cannot understand its material form until we understand the *laws underlying the form* – that is the *laws of the urban object itself*. Only through these laws can we understand the city as an object in all its social, cultural and psychological complexity. I do not contend that these laws of material form are *sufficient* for an understanding of the city; but I do contend that they are *necessary* for it. I do not hope to replace; social, economic and cultural theories of the city. But I do hope to complement them.

There are, I believe, three types of laws necessary to an analysis of the urban object:

- **type 1:** laws for the generation of the urban object itself, that is, laws governing the ways in which buildings can be aggregated to form towns or urban areas: these we might call the *laws of the object itself*;
- **type 2:** laws of how society uses and adapts the laws of the object to give spatial form to different types of social relation: these we might call the *laws from society to urban form*; and,
- **type 3:** laws of how urban form then has effects back on society – the old issue of architectural determinism, if you like: these we might call the *laws from urban form to society*.

If we accept the argument that architecture is made up of three disciplines, not one:

- construction (*firmitas*),
- style (*venustas*) and
- space (*commoditas*),

then I believe that construction has only the first type of law, style the first and second, while space, and only space, has all three types of law. In this paper, my argument is confined to space. I give examples of each type of law, and show how all three are involved in understanding urban form.

Spatial culture

As far as space is concerned, the three types of law, although analytically separable, eventually come down to one fundamental proposition: that human societies order their spatial milieu in order to construct a *spatial culture*, that is, a distinctive way of ordering space so as to produce and reproduce not *actual* social relations (the essential error of Modernist architectural determinism) but *the principles for ordering social relations*. Space is used sometimes to generate and sometimes to restrict the *field of encounter* of human beings and their symbols. How this happens depends on the forms of *social reproduction* involved. In all cases, however, space is not simply a *function* of the principles of the social reproduction: it is an *intrinsic* aspect of it, a necessary part of the social *morphology*.

Once we understand this critical relation, then we can begin to understand why it is that cities take on such different forms in different social and cultural conditions: why there are administrative cities and commercial cities, which are physically as well as sociologically distinct from each other; and why there are Islamic cities and Italian cities, each with their distinct characteristics as material forms and as cultural milieux.

The three types of law that underlie urban form

Examples of each of the three types of law are presented here in such a way as also to suggest a theory for the generation of urban form itself, focusing first on the origins of the urban block, then on its transformation in the creation of urban form itself, and finally its recent tragic disaggregation. In other words, I hope that my presentation will be read not only as an attempt to 'explain' urban form, but also as a radical critique of recent practice.

Type 1: Laws of the urban object itself

Through the structure of certain small settlements in the north of England (Figures 13.1a and 13.1b) or in the south of France (Figures 13.2a and 13.2b) we can discover the origin of the urban *block*. All four examples are of the kind that would normally be classified as 'irregular' or 'random', since they all lack the kinds of defining features that govern standard classifications: village greens, linear streets and so on. Yet, oddly, their very irregularity suggests that they might be – at some level – of a similar *organic* type.

Careful description can indeed reveal certain common properties. Each settlement is composed of dwellings that define, by their location and orientation, a continuous structure of open space which has a very characteristic form.

- First, there is constant variation in the width of the open space, so that 'fatter' areas of space are linked to other fatter areas through 'thinner' spaces, creating an effect rather like irregular beads on a string.
- Second, the structure of open space forms at least one ring, and maybe several intersecting rings, which means that there is always at least one alternative route from every point in the settlement to every other point.
- Third, although the maps do not show this, nearly every distinguishable segment of space, fat or thin, is directly adjacent to at least one building entrance, so that the spatial pattern seems to be in some sense defined by building entrances.

Taken together, these three properties bring to mind a primitive and irregular version of the urban block, with its outward-facing groups of contiguous buildings defining rings of space.

In spite of their individuality, all four settlements have in common what we might call, for want of a better term, a 'beady ring' structure: that is, a ring or rings of space formed like beads on a string by the orientation of dwelling entrances. What can explain this family likeness? None of the usual candidates – climate, topography, defence, land tenure systems and so on – seem to begin to explain the spatial peculiarities of this form. So perhaps we should explore what might seem, on the

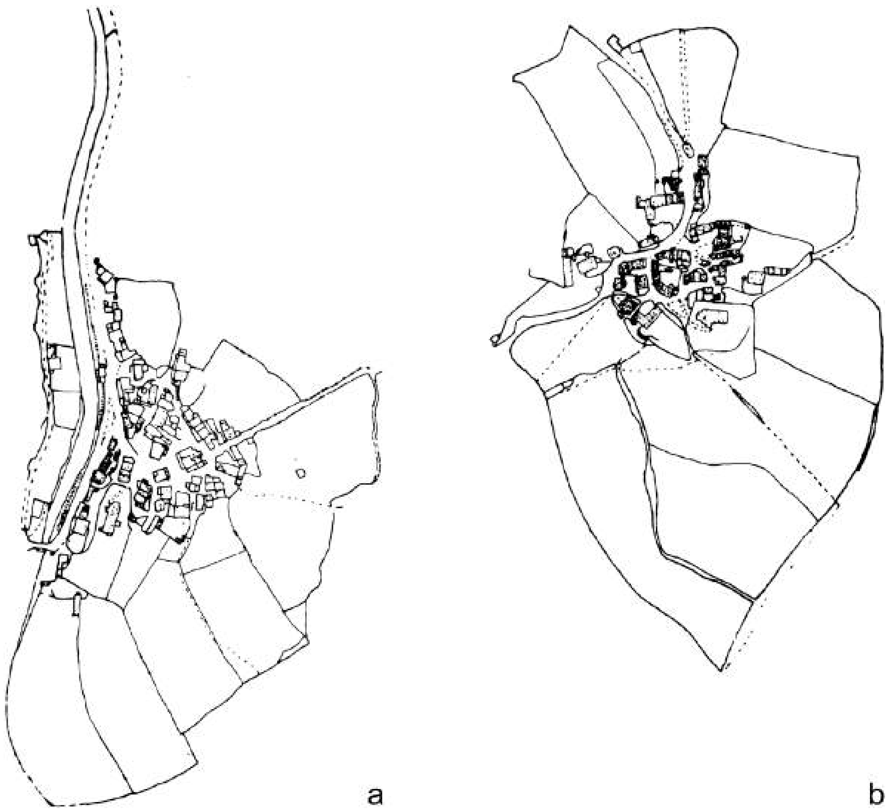


Figure 13.1 The structure of two small settlements in the north of England: (a) Muker (after OSC, 1888–1893); (b) Middlesmoor (after OSC, 1888–1893).

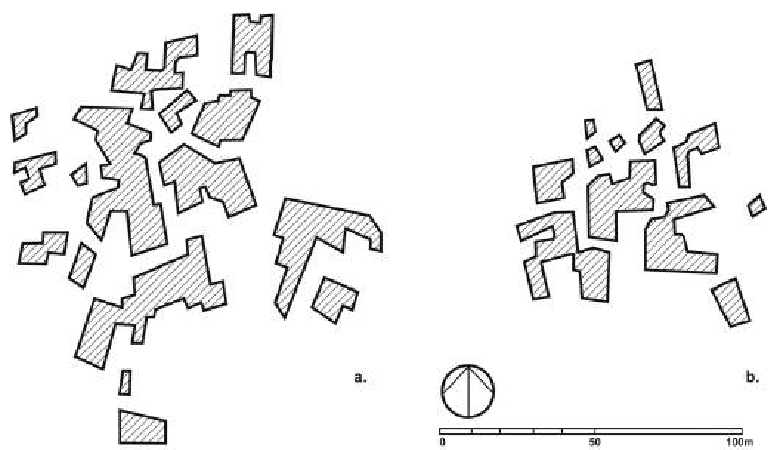


Figure 13.2 The structure of two small settlements in the south of France: (a) Perrotet; (b) Les Yves.

face of it, a more natural possibility: that the beady ring form arises from some process of *organic* growth.

A simple model can suggest such an aggregative process. Let us define an ‘elementary unit’ for this process as being made up of a simple square dwelling with one entrance and a segment of open space on the entrance side equal in size and shape to the dwelling (Figure 13.3). Now, let us define a process of random aggregation of such units, subject to two rules: each unit must, as it is added to the complex, be joined by its open space full facewise (that is, joining the whole side) to the open space of a unit already in the complex; and while full facewise joins of dwellings are allowed where they occur randomly, the joining of dwelling by their vertices is not allowed, since it is unlifelike.

Figure 13.4 shows a fairly typical result of such a process of aggregation, carried out by computer to ensure randomness, and frozen at about the size of a real ‘beady ring’ settlement. Figure 13.5 is a much bigger surface generated by the same process. It is clear that all three defining characteristics of the real beady ring structure – the formation of variable beads and strings of space, the formation of rings, and the adjacency of space segments to building entrances – are all realised in the simulated process. Of course, one might object that the form is over-regularised! But it need not be so. Figure 13.6 is a village layout generated on the computer by Paul Coates of UCL and Autographics Ltd in an effort to simulate some rather more linear beady ring forms that are found in the north of England. The beady ring process is, it seems, robust enough to survive most minor changes in the process. It is the accumulation of spatial relations that gives rise to the form.

What about the variations between the settlements in Figures 13.1 and 13.2? The main difference seems to be in the number and size of the rudimentary ‘blocks’. Those in Figure 13.1 have several small blocks, while those in Figure 13.2 have one dominant block surrounded by several small ones. This difference in fact could be modelled by a simple change in the generative rules. In the process already described the joining of two dwellings – as opposed to their spaces – occurred only where the process generated it randomly. But we can make this subject to a rule, one perhaps which increases the *probability* of dwelling joins. The effect of this will be, as we might imagine, to increase the size and decrease the number of the blocks as we increase this probability and vice versa; and this of course is the main difference between the Figure 13.1 type and the Figure 13.2 type.

We can then extend this change in the process to its limit until we require the dwelling to join and allow space joins only where they occur randomly. This will generate a single irregular block, penetrated by deep, wandering three-sided courtyards (Figure 13.7). Such a phenomenon can often be found in settlements, especially where some external restriction – say, the presence of a highly restrictive town wall – forces dwellings into greater contiguity. Could some of the block shapes in Figure 13.8 perhaps be influenced by such factors?

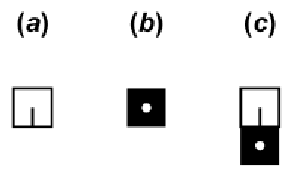


Figure 13.3 The elementary units used as inputs in the simulation model of organic growth: (a) notional dwelling with entrance; (b) open space; (c) elementary unit with open space on the entrance side.

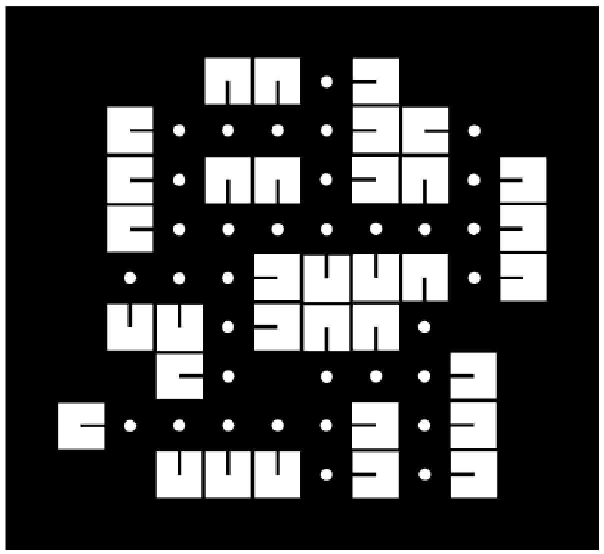


Figure 13.4 Computer-generated beady ring surface – simulation of organic settlement growth.

Other variations follow from changing the *scope* of the rules. In Figure 13.9, for example, the basic beady ring process seems to be present, but to have become more regularised so that lines of sight are much more extended than they would be in a randomly generated aggregation. Such an effect can be produced by a rule that requires each added unit to take into account not only how it relates to an immediate neighbour, but also how it relates to other dwellings some distance away, that is, the *scope* of rules is extended. It was in fact once suggested to me by a local builder that new houses could once be located at will, provided they did not obscure sight lines from any dwelling entrance to any other. Such a rule would be likely to have the effect, in an otherwise random process, of producing a greater tendency to linearity of the kind seen in Figure 13.9.

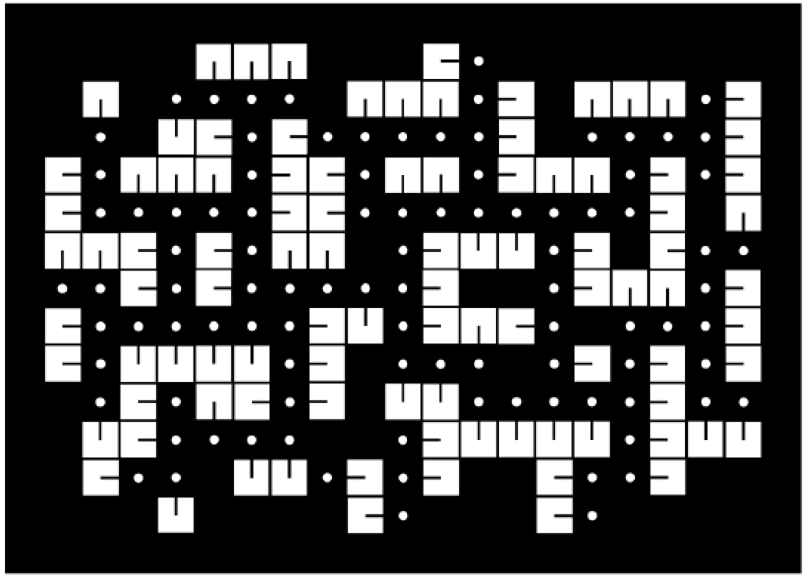


Figure 13.5 Larger computer-generated beady ring surface.

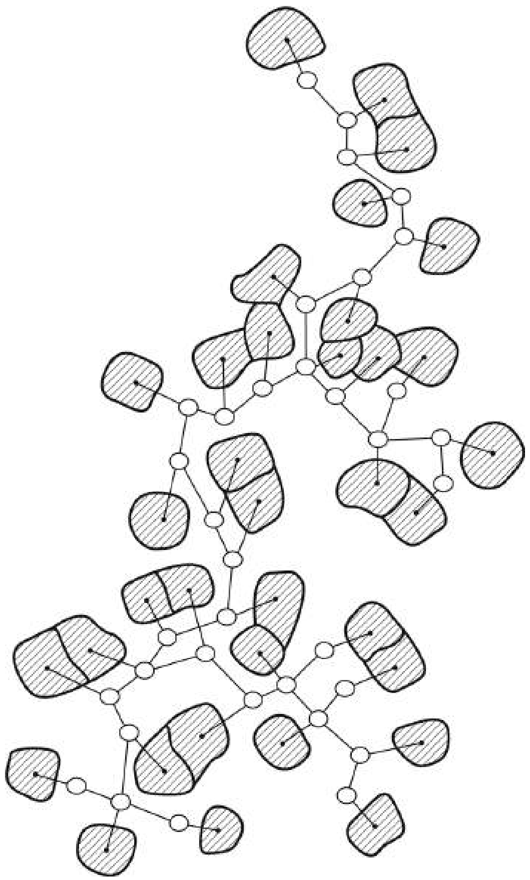


Figure 13.6 'De-geometrised' computer-generated village layout.

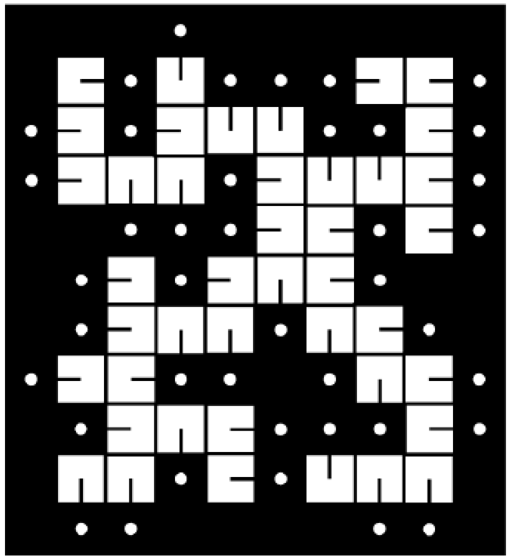


Figure 13.7 Computer-generated obverse beady ring surface.

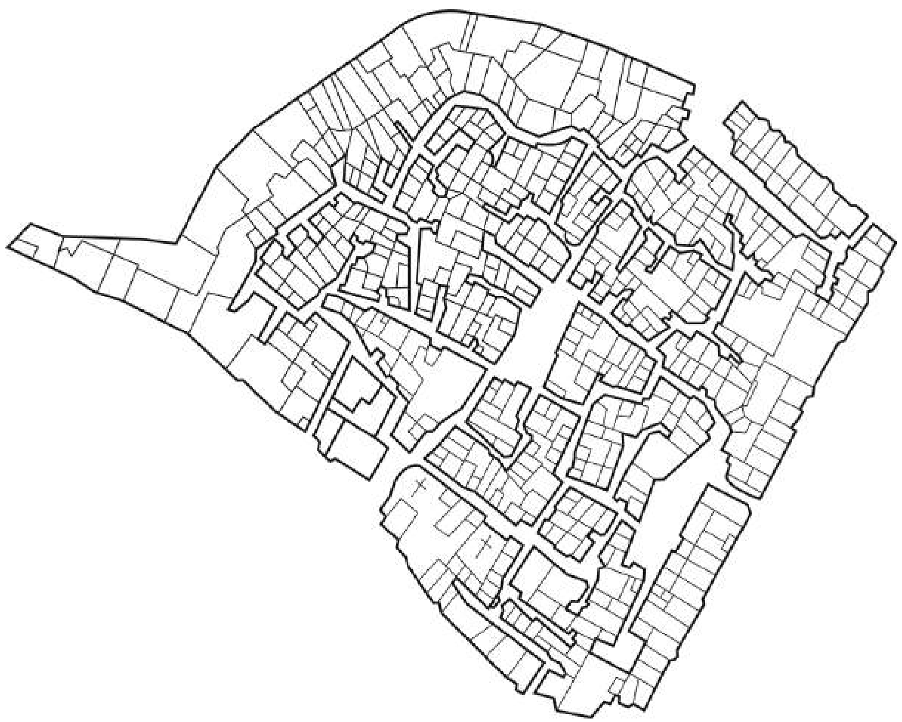


Figure 13.8 The walled town of Cisternino in southern Italy.

Figure 13.9 Layout of the small town of Gassin in the Var region of France.

Can we then say that the various processes 'explain' the forms? The question is difficult to answer on the basis of the completed forms, since we need to consider the historical data on the step-by-step growth of each settlement. Such data are, unfortunately, hardly ever obtainable. We have, however, tried to test the proposition in the case of the beady ring settlements in the south of France by mapping some 77 apparently random clusters of buildings in the vicinity of the two illustrated settlements. Figures 13.10a and 13.10b are examples of small clusters, Figure 13.11 a larger one (drawn to a slightly smaller scale). Analysis of the sample shows:

- first, that every cluster of sufficient size takes on a clear beady ring form, subject to topographical constraints; and,
- second, that all smaller clumps are either compatible with a beady ring process, or strongly suggestive of it. Figure 13.10a is a 'compatible' case and 13.10b a 'suggestive' one.

Suppose, then, that the beady ring process were involved in the formation of these settlements. What would this tell us, philosophically? Surely it would tell us that to explain settlement forms we need two kinds of knowledge, not one. We need knowledge of real historical events and the socio-economic processes in which they were embedded. But this, on its own, would never explain the morphological regularities. To explain these, we also need to know that the local rule of 'open space joins' applied to an otherwise random process of growth will itself lead to the beady ring form. The historical process has, as it were, activated a morphological process.

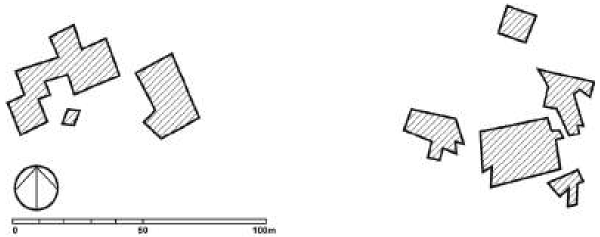


Figure 13.10 Small clusters of buildings in Esquerade (left), Les Bellots (right) in the south of France.

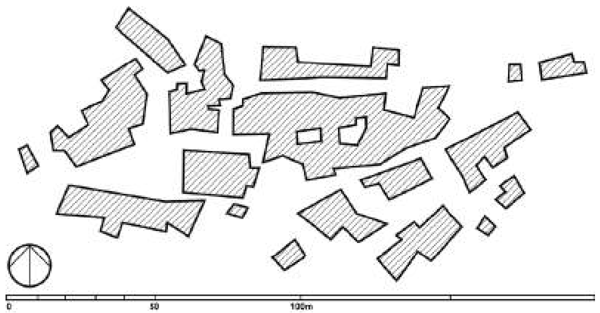


Figure 13.11 Larger cluster of buildings exhibiting a clear beady ring form – Les Petits Cléments in the south of France.

But the morphological process is *lawful in itself*, and it is this internal lawfulness that in the last analysis actually produces the global morphological regularities. In other words, knowledge of laws of type 1, the laws for the formation of the object itself, is indispensable for an understanding of the material form. Moreover, it seems highly likely that the best key to the 'spatial culture' of the settlements might be the generative rules underlying the material form, rather than simply the form itself. We can explore this possibility further by considering some rather more complex cases, and looking for laws of type 2.

Type 2: Laws from society to space

Let us compare a representation of the town of Gassin (Figure 13.9) again, with the open space structure rather than the blocks in black (Figure 13.12), with a similar representation of another southern French town, Apt (Figure 13.13). Gassin is a small town on the ridge of a hill. Apt is a large town on a flat plain by a river. I argue that in spite of these locational and topographical differences, these two towns share a common morphological *genotype*, one which compels a *social* interpretation of *spatial* form. In explaining this, I also explain something of the method of spatial

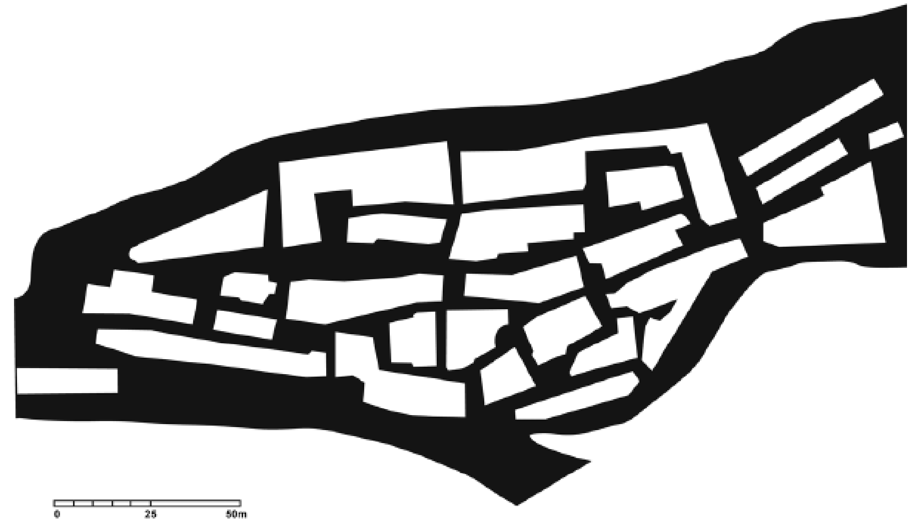


Figure 13.12 The open space structure of Gassin in the south of France shown in black with buildings shown in white.

Figure 13.13 The open space structure of Apt, Vaucluse, in the south of France shown in black with buildings shown in white.

analysis we have developed called 'space syntax', and use it to show that the more urban type of order found in these towns arises, to a considerable extent, from the transformation of the urban block – that is, the rules of aggregation are assigned to the block rather than the individual building.

What has to be explained in these cases is the partly regularised but still irregular pattern of open space structure, and how it is produced by the arrangement of buildings. We can begin by observing that, although irregular, both settlements are composed of blocks surrounded by space which form intersecting rings, and therefore take a form which has a topological resemblance at least to an urban grid. We might call them '*deformed grids*' – and note in passing that the great majority of towns constructed by human beings throughout history are based on some variant of the deformed grid.

In what sense is the grid deformed? I suggest that it is simpler than it looks and that the beady ring settlements have already given us the basic concepts we need. All that is needed is to make them more rigorous.

First, the grid is, I suggest, deformed *two-dimensionally* to form a set of identifiable *convex* spaces, some fat and some thin. Figure 13.14 is a break-up of the open space of Gassin into its fattest convex spaces (a piece of space always forms part of the fattest space it can). This immediately shows that the principle noted in the beady rings, that nearly *all* identifiable spaces, thin as well as fat, have entrances opening onto them, is maintained in Gassin. The same is also true of Apt. The adjacency of convexly defined spaces to entrances must therefore be held to be a *genotypical* spatial property of these settlements.

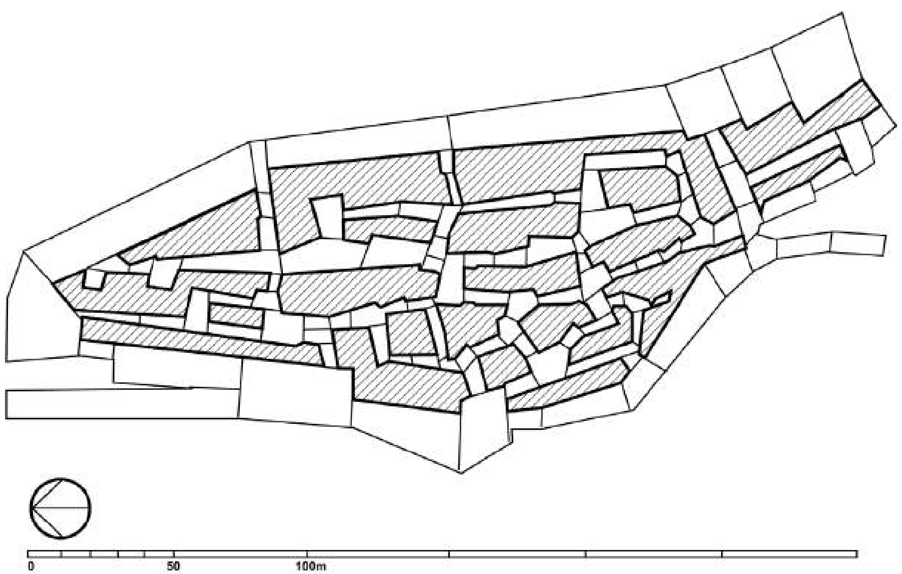


Figure 13.14 Convex map of Gassin in the south of France.

But it is more than a purely spatial property. A convex space is one in which each point is directly visible and accessible from every other point. If building entrances are systematically related to convex spaces, then it must be because they are meant to participate in this 'all-play-all' mutual surveillance. The 'convex space entrance' principle ensures that virtually all, not just some, spaces in the settlement are, in a sense, under the control of entrances and, potentially, of people who may come and go through them. We are therefore dealing with a genotypical *socio-spatial* property of the settlements.

Even more striking socio-spatial properties appear when we consider the second kind of deformation of the grid: the *one-dimensional* or *axial* deformation. Figure 13.15 is an 'axial map' of Gassin, formed by drawing the longest and fewest lines of direct access and visibility that cover all the convex spaces of the settlement. Figure 13.16 is a similar map for Apt. Certain properties are immediately apparent. First, although there are far more lines in the settlements than would be needed to cover a perfectly regular grid (how many more is one index of how axially deformed the grid is), there are not so many as to suggest terms like 'labyrinthine.' The two-dimensional deformity does not always entail one-dimensional deformity. On the contrary, it is a striking feature of both settlements that axial lines seem to run through whole series of convex spaces.

This is an important property in defining 'urban' spatial experience. If a convex space is that region around us where all points are visible and directly accessible from all other points, then axial lines tell us about *some* points in other, perhaps remote, convex spaces which are also visible and directly accessible to us. Through this relation between convexity and axiality in space, we are in effect

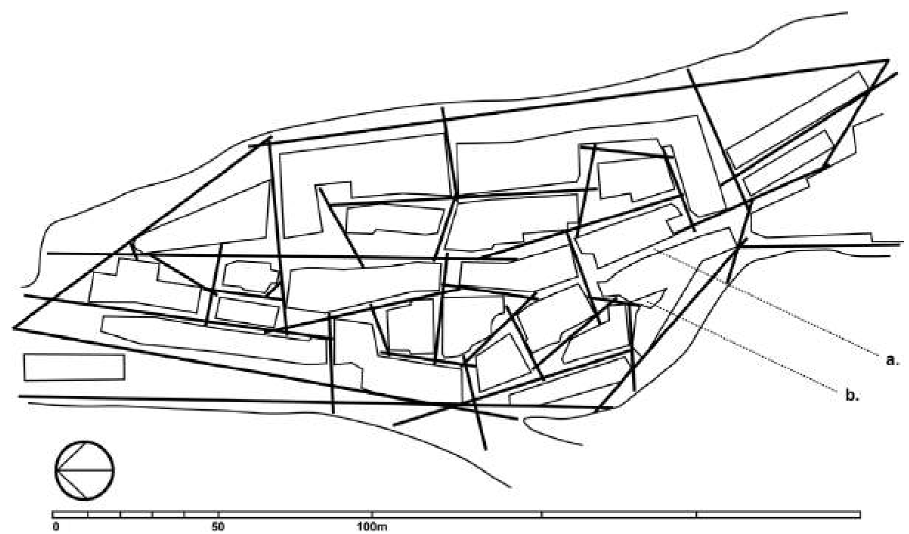


Figure 13.15 Axial map of Gassin in the south of France. See Figure 13.17 for explication of 'a' and 'b'.

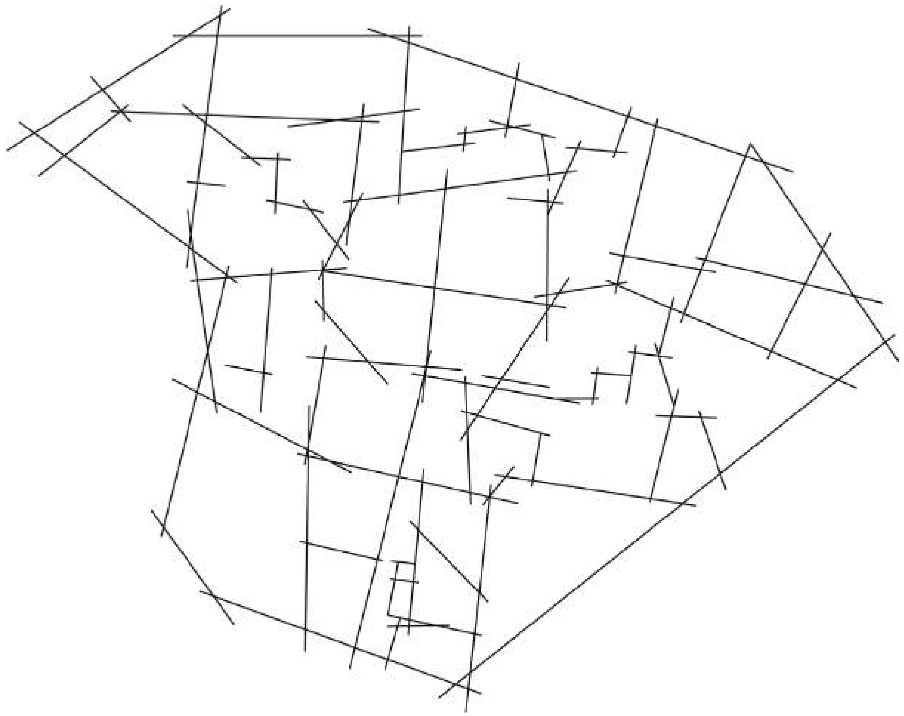


Figure 13.16 Axial map of Apt in the south of France.

given two kinds of information from space: complete *local* information about the space we *are in* through the convex organisation; and partial *global* information about spaces we *might go to* through the axial organisation. In urban space, we are in effect given information about *two scales at once*. This *compression of scales*, to my mind, comes close to being the essence of urban spatial experience.

But to arrive at the full picture of the *socio-spatial genotype* underlying these two settlements, we need to go beyond direct experience and consider the spatial pattern of the settlement *as a whole*. Because it is the most *global* representation of the spatial structure, we will consider the axial map on its own, and introduce some of the numerical techniques of space syntax analysis.

Of these, the most important is a measure we call *integration*. If we look at the axial map, we can see that every other line is linked to every other line, either directly or by way of a certain minimum number of intervening line segments, or 'steps'. We will call this the property of *depth*. A line is as deep from another line as the least number of steps that have to be used to go from one to the other. Thus, a space is at *depth 1* from another space if it is directly connected to it, at *depth 2* if there is one intervening step, and so on.

It follows that every line has a certain depth from every other line. The *integration value* of a line is a mathematical way of expressing the depth of that line from *all* other lines in the system (Hillier and Hanson 1984, for the equation). It is not at all obvious that these values will differ significantly from one line to the next; but that they do is one of the most significant properties of architectural and urban spatial configurations.

A graphical way of expressing that difference is shown in Figures 13.17-top and 13.17-bottom. Figure 13.17-top is a 'justified graph' of the system of lines as seen from the point of view of the long line (marked 'a' in Figure 13.15) descending from the edge of the settlement at half past two towards the centre, the most 'integrating' line in the settlement. The points in the graph represent the lines, and the connections represent the intersections of lines. The long line, 'a', of Figure 13.15 is the 'root' of the graph, and each level of 'depth' away from that line is aligned vertically, so that the height of the graph shows how integrated the line is: the shallower, the more integrated, and vice versa. Figure 13.17-bottom is a similar graph drawn from a short line (marked 'b' in Figure 13.15) in the bottom

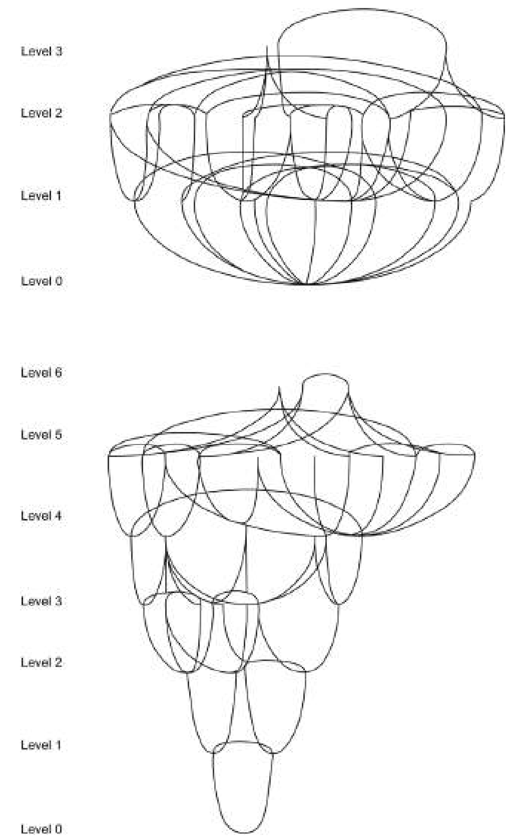


Figure 13.17 A justified graph of spatial relations in Gassin (top) from an integrated line; (bottom) from a segregated line.

right complex of Figure 13.15. The greater height of the graph shows that the system has more 'depth' from that line, and that line is therefore less 'integrated', or more 'segregated'.

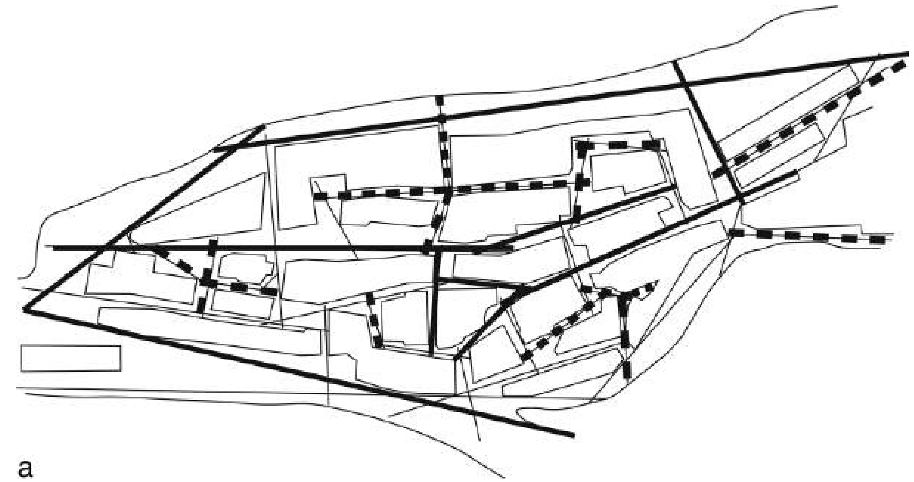
The *integration core* of the settlement is then the 10 per cent (or 5 per cent or 25 per cent, depending on what is to be shown) of most integrating lines, numbered in order of integration. Figure 13.18a is the integration core of Gassin and Figure 13.18b that of Apt, marked in black, with the dotted lines representing the most segregated lines. Both cores take the form we call a *deformed wheel*: a semi-grid or 'hub' of lines in the interior of the settlement is linked by lines, or 'spokes', in several directions to peripheral or 'rim' lines. The segregated lines then cluster in the interstices formed by this wheel structure. Underlying the differences in size, location and topography, therefore, the two settlements share the same '*deep structure*' or *genotype*.

Is this, then, also a *socio-spatial* genotype? The form of the core surely gives a clue. By linking the interior of the settlement to the periphery in several directions – and always in the direction of the main entrances to the settlement and the neighbouring towns – the effect of the integrated lines is to access the central areas of the town from outside, while at the same time keeping the core lines close to the segregated areas, in effect linking them together. Since the core lines are those that are most used by people, and also those on which most space-dependent facilities like shops are located, and the segregated areas are primarily residential, the effect of the core is to structure the path of strangers through the settlement, while at the same time keeping them in a close interface with inhabitants moving about inside the town. The structure of the core not only accesses strangers into the interior of the town but also ensures that they are in a constant *probabilistic interface* with moving inhabitants. Indeed, it seems reasonable to propose that the spatial structure of the settlement exists *in order to* construct this interface.

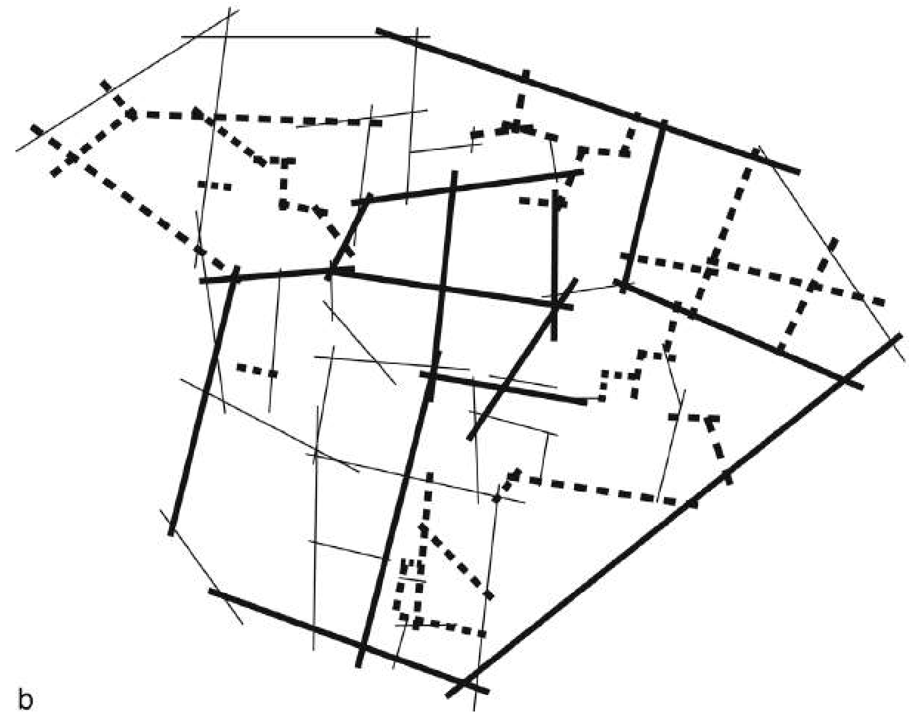
Not all towns have this type of structure: it depends on what kind of interface is to be constructed. In the two settlements we have just looked at, the structure of space has to do with the practicalities of moving people into and around the settlement. Space plays a largely *instrumental* role. In other types of town, however, space appears to be organised more to relate *buildings* of *symbolic* importance, something that is largely lacking in the towns we have looked at.

Let us begin with an extreme example: the pre-Columbian American 'town' of Teotihuacan (Figure 13.19). In spite of the fact that the compounds seem to be based on a standardised metric, the town seems to lack any obvious form of axial logic – apart from, that is, the single central axis which passes between the Great Compound and the Citadel, crosses the town and strikes full face onto the Pyramid of the Moon at its other extremity. In fact, the more we look at it, Teotihuacan seems to contradict the spatial logic of the 'instrumental' towns at every point. In spite of its greater geometricity, the open space is more broken up, both convexly and axially.

There is no consistent relation between convex spaces and building entrances. On the main axis, there are very few, if any, everyday building entrances, and to



a



b

Figure 13.18 (a) Axial map of Gassin with the 25 per cent integration core shown in heavy black and the 25 per cent most segregated lines shown dotted; (b) axial map of Apt with the 10 per cent integration core shown in heavy black and the 50 per cent most segregated lines shown dotted.

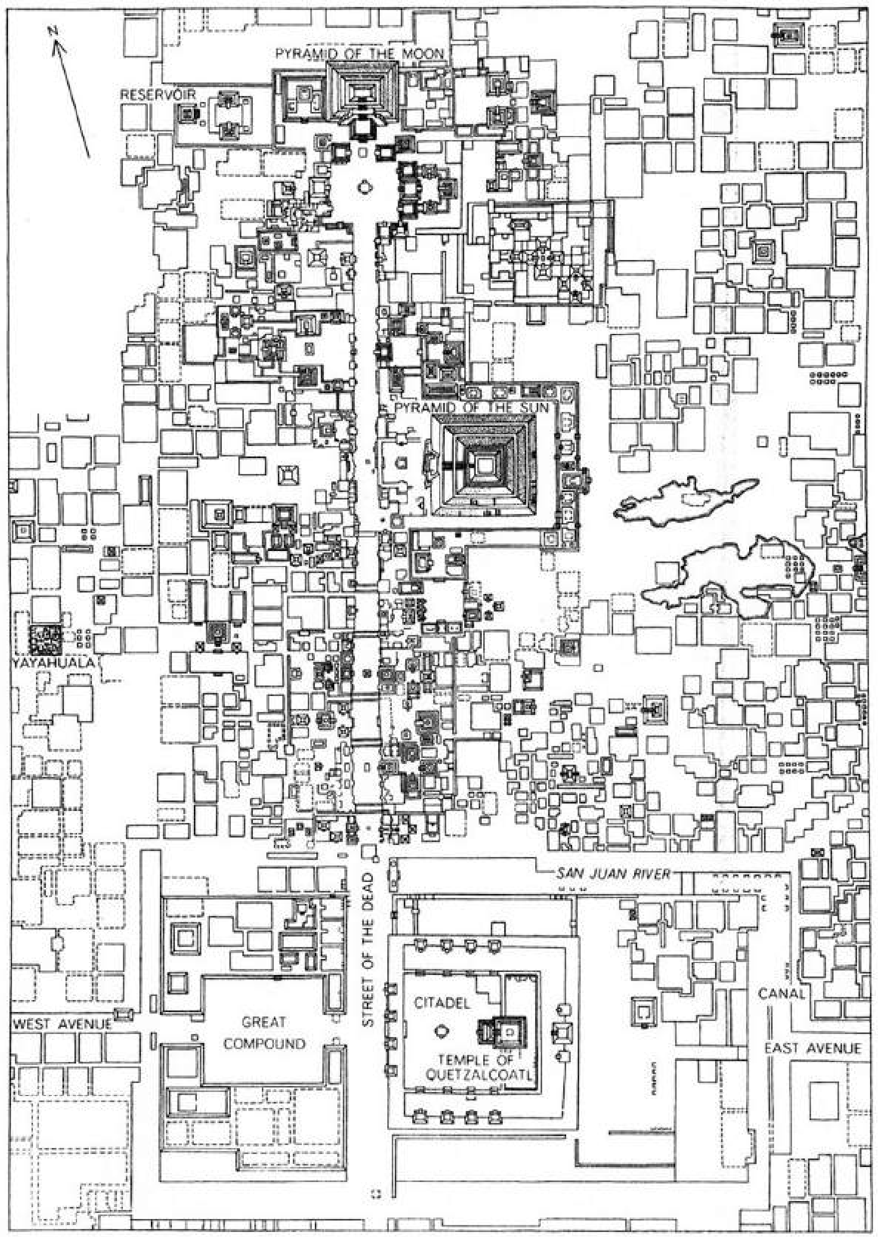


Figure 13.19 Teotihuacan – general map of the central zone of the city, showing cruciform plan and location of the principal constructions.

a great extent the axis is bounded only by ceremonial structures. This axis also approximates in itself to a convex space. Instead of an axial line penetrating many convex spaces, the convex space is expanded to become coterminous with a single axial line. Nor does this axis link edge to centre in the manner of the previous cases. It runs between two major buildings and strikes another full face. This orthogonal striking of major axial lines on building facades is again a property not found in the previous two cases. In these, most axial lines 'glance off' major and minor buildings indiscriminately at open angles, suggesting movement beyond the point where the line strikes the building. Only in the more segregated areas do lines strike buildings more orthogonally.

I suggest that this basic morphological scheme is the one that tends to be used in towns which are less concerned with the *production* of everyday life and more with the formal *reproduction* of social structures. The orthogonal end-stopping of *axes* on major buildings, in contrast to the way in which lines glance off buildings in instrumental settlements; the expansion of convexity to cover most of the length of major axial lines, as opposed to the linking of convex spaces by axial lines; the removal of everyday buildings from selected major axes, as opposed to their ubiquitous presence; the running of major axes from major building to major building, rather than to direct movement from outside to interior and out again; the strong contrast between a few very powerful spaces and the much more even distribution of integration throughout the settlement – all these are the means by which the spatial structure of a town is changed from being an *instrument of pervasive but variable co-presence* to a *symbolically ordered ideological landscape* expressive of the forms of power in a society. Versailles and Brasília (Figures 13.20 and 13.21) are examples of such structures.[9] These two types of town (there are also many others, of course) are not simply unlike each other: they are *systematically* unlike each other. The principles they use are in key morphological senses the opposite of each other; and they are so because they realise in space quite different social schemes and priorities. And their encounter consequences are different. In instrumental towns, space is differentially used but is everywhere dense. In symbolic towns, many important spaces are much more sparsely used. The spatial genotype, as it were, freezing the instrumental use of space, thus emphasising its primarily symbolic nature.

In many towns, of course, one can find both types of spatial ordering, but, I suggest, always in such a way that the functions of everyday production and exchange are realised through the first, instrumental principles of spatial patterning, while those that have to do with the functions of social reproduction are realised through the second, or symbolic principles of order. In other words, the means by which social formations impress themselves on urban form are themselves subject to regularities which, in turn, are the product of underlying laws. To understand towns, therefore, we need an understanding also of this second type of law, and these could not have been understood without prior knowledge of laws of type 1.

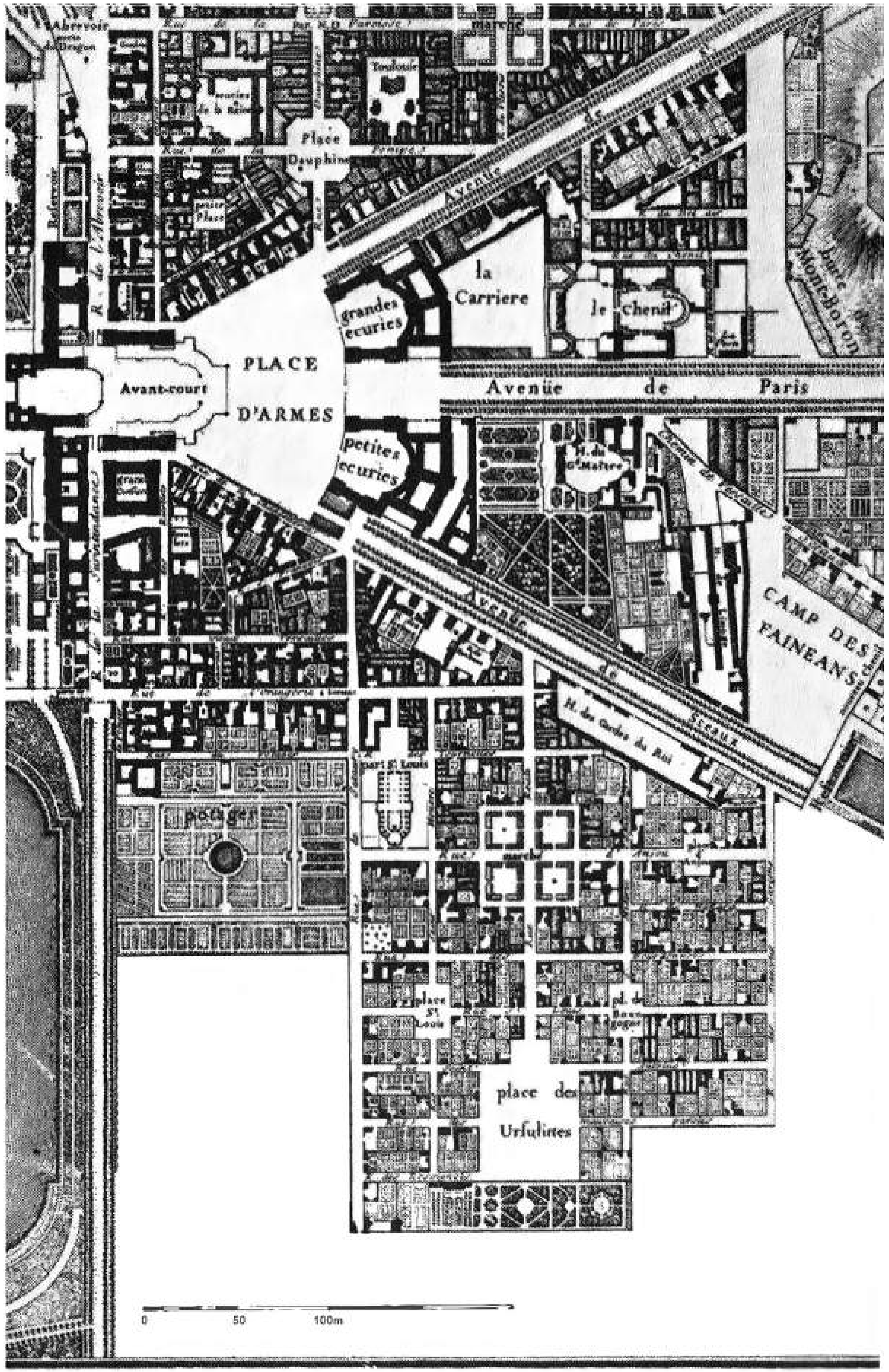


Figure 13.20 Plan of Versailles – axiality used to create a symbolic landscape.

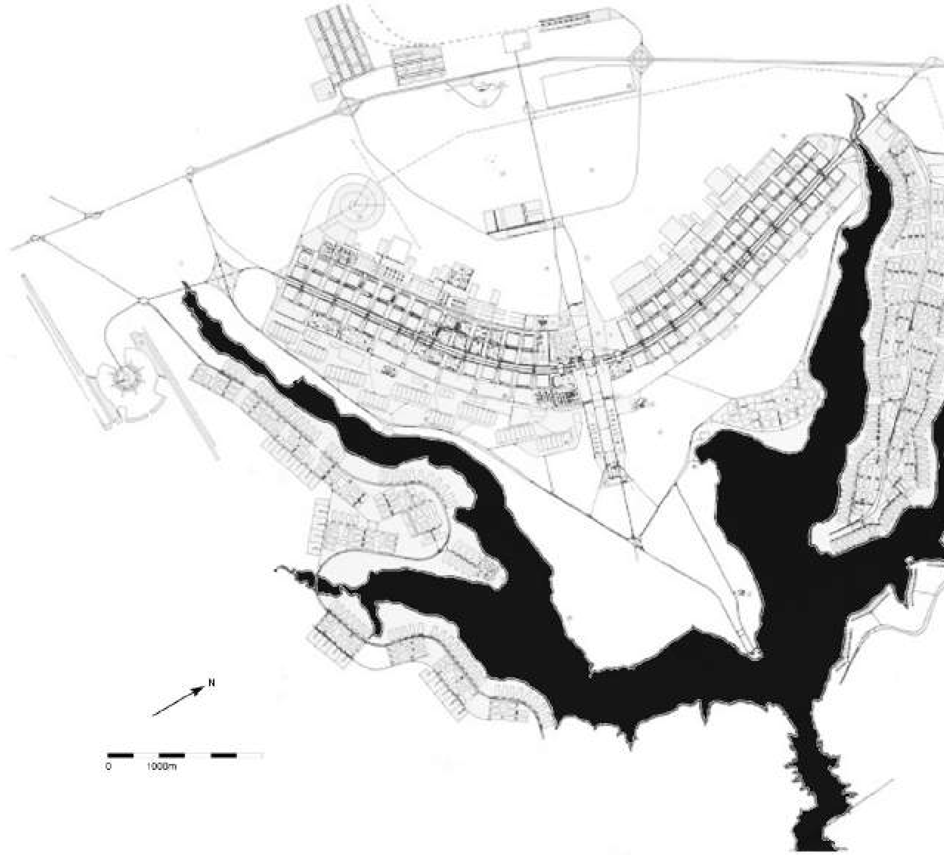


Figure 13.21 Plan of Brasília – axiality used to create a symbolic landscape.

Type 3: Laws from space to society

Laws from space to society are those by which determinable and describable effects on *people* arise from spatial form. This is the dangerous ground of architectural determinism, a hypothesis which I, among others, have argued against in its naïve forms. I must therefore state my case with great care. I argue that the belief that spatial form has *no* effects on people and society is patently absurd. If this were the case, then we could design every monstrosity without penalty. My proposal is that the determinable effects of spatial form on people are both limited and precise. Spatial form, I argue, creates the *field of probable* – though *not* all possible – *encounter and co-presence within which we live and move*; and whether or not it leads to social interaction, this field is *in itself an important sociological and psychological resource*. I will try to show that this field has a definite *structure*, as well as properties of *density* or *sparsity*. It therefore deserves a name. I will call it the *virtual community*, meaning that it exists even though it is latent and unrealised. The virtual community is the *direct product of spatial design*.

While I believe I can demonstrate the *existence* of the virtual community as an entity with a definite structure, I can only express my *belief* that it is a sociologically important entity, and only *argue* that the relationships that have usually been advanced in support of architectural determinism – for example the relation between architecture and crime, or that between architecture and community formation – if they exist at all, exist through *the intermediary role of the relationship between spatial form and the virtual community*. Without knowledge of this fundamental relation, it is impossible to identify *any* systematic relation between architecture and people. Because the virtual community is the product of *spatial configuration*, it has only been possible to detect its existence by the use of space syntax analysis, coupled to very precise and systematic observation of where people are in space and how they move. Systematic observation of spaces on the basis of a syntactic analysis shows that the rates at which people use space and move through it are statistically reliable properties of spaces, and can be assigned to spaces as *encounter rates* for those spaces. Different encounter rates can be established for different categories of people: moving and static people, men and women, adults and children, and so on. Because they are numbers, these encounter rates can then be correlated with the values (for example, integration values) assigned to the spaces by syntactic analysis. The pattern of correlation will then allow us to build up a picture of the fundamental relationship between the *spatial configuration* and *encounter pattern* of an area. This relationship is the *structure* of the virtual community for that area. It turns out that this relationship is, again, in itself, *lawful*. The laws of the socio-spatial entity we call the virtual community are then my laws of type 3: laws from space to society.

A considerable body of research now exists on the relationship between the spatial syntax of an area and its encounter pattern.[10] Here I simply illustrate this work by reference to a single large-scale study in North London. Figure 13.22 is the Ordnance Survey map of a section of North London stretching from King's Cross in the west, the limits of the City in the east, and Holloway and Canonbury in the north. Six local studies were carried out in the area: two of nineteenth-century street systems, three of housing estates built in the last quarter century and one a mixture of estates and a residual street system.

Figure 13.23 is an axial map of the area considered as a street system only. Figure 13.24 is the same axial map with the ground level of the estates added. The contrast is instructive. Although the estates cover only a fraction of the total surface, they more than double the number of lines in the system, with an average length of line much shorter than in the street system. This is true of *all* the estates, even though some are early modern slab blocks, others are 1960s blocks-in-courts megastructures, and other recent neo-vernacular attempts to copy the urban or village past. *All*, it turns out, dramatically *reduce* the scale of effective space and render it much more *localised* and *deep* from the street system, and therefore substantially more segregated from the urban fabric as a whole. This can be

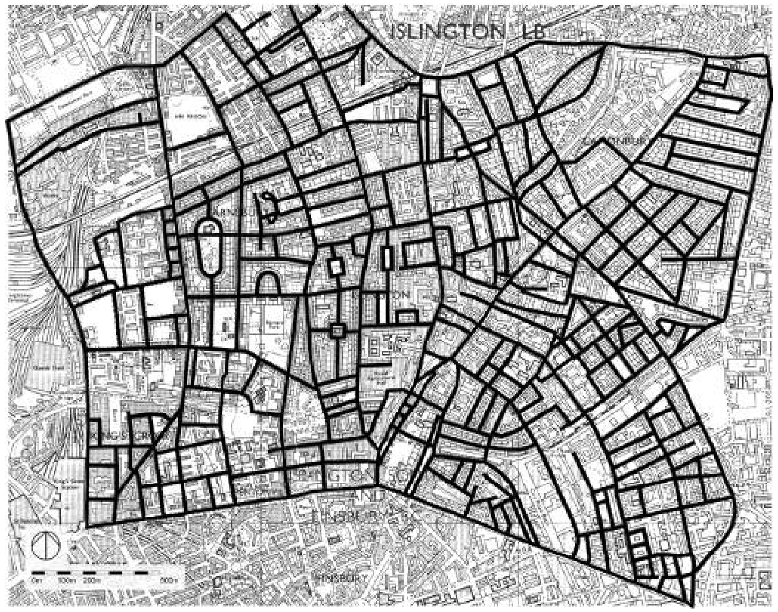


Figure 13.22 Map of larger area of study – Islington, North London. National Grid, circa 1979. Updated: 30 November 2010 using EDINA Historic Digimap Service.

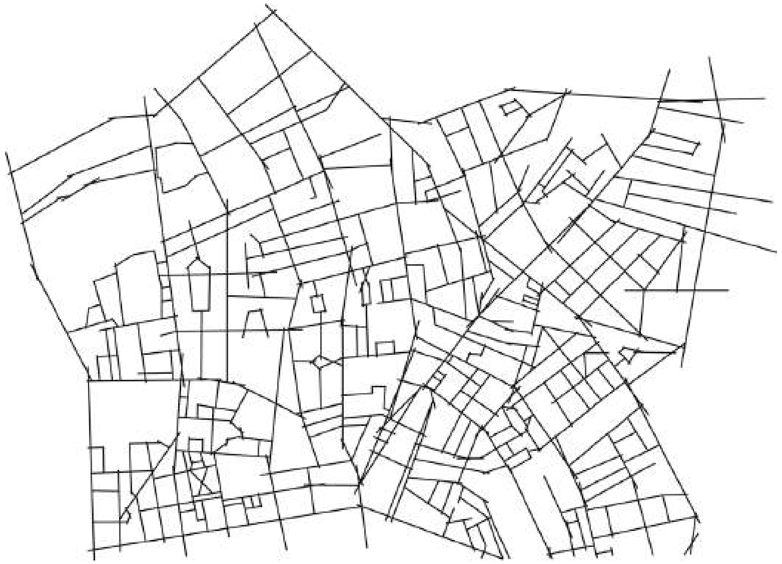


Figure 13.23 Axial map of Islington (North London) excluding estates.

expressed mathematically: the *mean integration* of lines in the street system is .76, a typical urban value for London, while that of the estates is .96 – the higher number expressing substantially more depth and less integration.

But the street areas and estates also differ on another fundamental mathematical property, which we call *intelligibility*. Intelligibility is the degree of correlation that exists in a system of spaces between how *connected* each line is to others, that is, how many immediate neighbours it has intersecting it, and how *integrated* it is into the system as a whole. We call this measure *intelligibility* because the number of lines intersecting a line can be directly seen from that line, while the property of integration cannot be directly seen since it expresses the depth of a line from every other space in the system, most of which will be invisible. The degree of correlation between *connectivity* and *integration* for lines in a system therefore expresses the degree to which the *local* and *visible* properties of spaces are a good guide to the *global* and *merely inferable* position of spaces in the area as a whole. In the case of Figure 13.24, the degree of *intelligibility* in the street areas is .61 (a high value expresses *more* intelligibility), while for the estates it is only .26. This mathematically captures the intuitive sense that the spatial structures of the estates somehow lack structure and intelligibility, even where they are relatively geometrical.

Now let us look at some of the results of encounter studies. Figure 13.25 is the map of one of the street areas, the St Peter's Street area, Figure 13.26 is its black-on-white space structure, and Figure 13.27 is its axial map with integration core and segregated areas marked in the usual way. The most integrated lines are

Figure 13.24 Axial map of Islington (North London) with estates included.

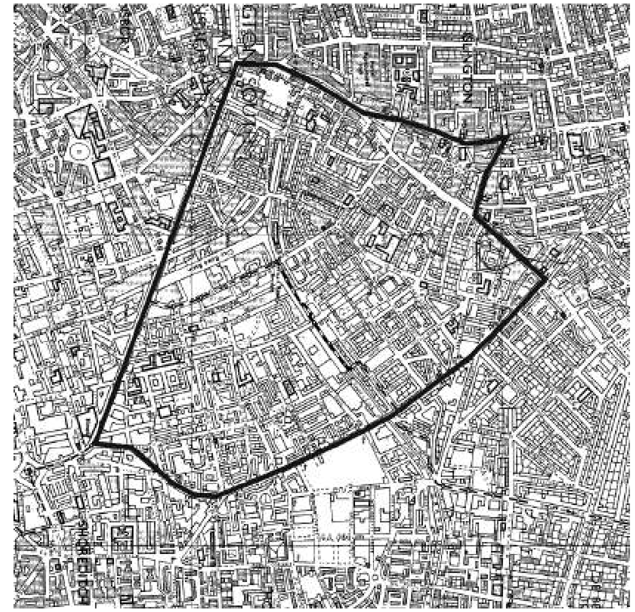


Figure 13.25 Map of the St Peter's Street area in Islington, North London. National Grid, circa 1979. Historic, using EDINA Historic Digimap Service.

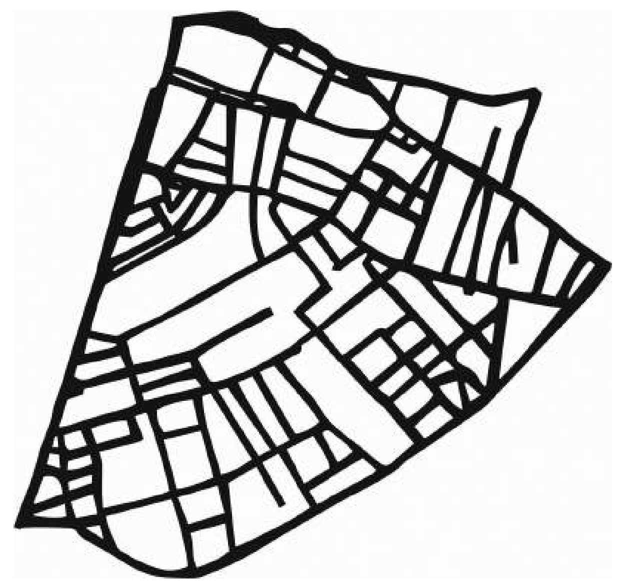


Figure 13.26 St Peter's Street area in Islington, North London, with the space structure shown in black and the buildings shown in white.

Figure 13.27 Integration core of St Peter's Street area in Islington, North London.

the relatively short lines that pass through the village centre for the area, with shops and various other facilities. The next most integrating lines then link this 'village centre' to the most important peripheral lines of the area's 'supergrid', which also act as part of the area core. In effect, the area reproduces something like the 'deformed wheel' core that we noted in the two original settlements. This, it turns out, is true of most inner London named areas, like Soho or Covent Garden.

Figure 13.28a is then a scattergram in which the vertical axis is the *integration value* of lines and the horizontal axis the mean observed *encounter rate* for moving people on those lines. If this relationship were perfect – that is, if integration were a perfect predictor of encounter rate – then the points would lie in a precise line at 45 degrees. As it is, the distribution gives a powerful approximation of this, with a statistical correlation of .7742 (a perfect correlation would be 1 and a random relation 0). This is about average for an urban area. Other spatial parameters (for example, the simple connectivity of lines) will also show a good correlation, but in most types of urban area, integration is the best, and can be used with considerable reliability as a means of predicting the general movement pattern of a layout at the design stage.

However, several other important spatial effects can be discovered by pursuing the correlation between integration and encounter rates for moving people.

- First, in Figure 13.28a, the scattergram is rather lumpy. It turns out that if we divide the area into localised groups of linked lines – while still calculating integration in terms of the whole area shown in Figure 13.27 – then the correlation between integration and movement within these groups is even stronger, usually at about the .9 level. This does not always happen in an urban area. It only happens where there is some tendency for identifiable subareas to form within the main area. In this case, the subareas are formed by the divisive effects of a canal, an industrial development and housing estates, one very large. Part of the lumpiness of the scattergram reflects this tendency towards subareas. It is important.
- Second, the integration values used in Figure 13.28a are calculated only on the basis of the street system, ignoring the ground-level interiors of housing estates. If integration values are re-calculated to include these, then scattergram 13.28b is the result: a marked deterioration in the correlation. This shows in a very simple way that the interior space patterns of the estates are unrelated to the way in which the street layout is producing a predictable encounter pattern. The predictive power lies only in the street pattern itself.
- Third, if we then re-calculated the integration values to take into account the much larger area shown in Figure 13.22 (or indeed any larger area than the one shown in Figure 13.27), then once again there is a dramatic deterioration in the correlation. This shows, again, that the encounter pattern of the area is very much a product of the area itself, and that it is not much affected by the pattern of other areas. The St Peter's Street area thus operates to some extent as a natural area within the larger urban framework, to which it is, however, highly connected and with which it shares the all-important supergrid. Optimising correlations between encounter rates and integration by varying the scale of the area referred to in the integration calculation thus offers a powerful method for identifying 'natural areas' within a larger urban system.

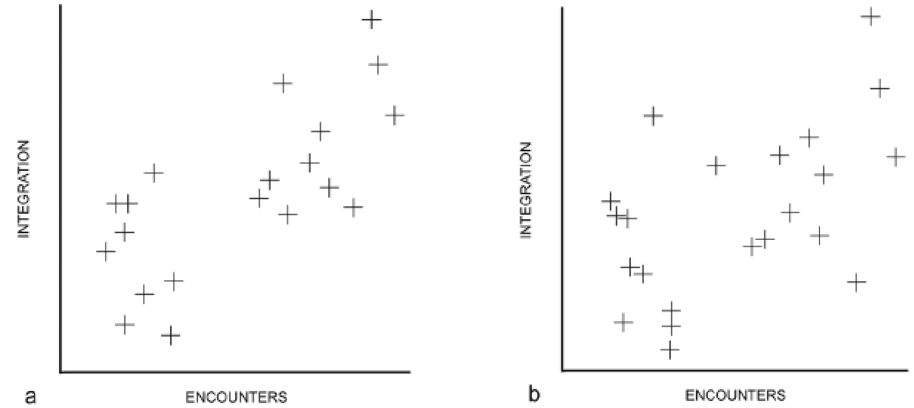


Figure 13.28 (a) St Peter's Street urban area; correlation of encounters with integration (r = 0.7742); (b) with estates included; correlation of encounters with integration (r = 0.5983).

Whatever reference area we use to measure it, *integration* is nearly always the best spatial parameter for predicting the *encounter rates* for moving people. People construct their patterns of movement, it seems, according to the picture they have of the *axial depth* of spaces from each other, with reference to a fairly large system of spaces. Effective spatial knowledge is thus at some level global knowledge. It is this that gives rise to one of the most fundamental of all spatial properties of urban areas: that the knowledge people have of the spatial structure of an area is also knowledge of encounter potential. Our picture of the area mixes space and people, because encounter rates are *predictable* from spatial structure, and spatial structures are arranged to be so predictable.

There is, however, a significant caveat. This relationship – the very basis, I would suggest, of the virtual community – only exists to the extent that the area is intelligible (in the mathematical sense in which we have defined it). Once again, this can be demonstrated.

In the estates shown in Figure 13.24, *intelligibility* breaks down, whether the estates are considered on their own or as part of some larger area. In close parallel, the *predictability* of encounter rates from spatial structure within the estates also collapses. Figure 13.29 is a scattergram of the relationship between *intelligibility* (the correlation between connectivity and integration) and *predictability* (the correlation between *integration* and observed *encounter rates* for moving people) for the six studied areas. As *intelligibility* increases, so *predictability* improves almost linearly.

But it is not only the predictability of encounter rates from integration that falls in the estates. The encounter rates themselves fall to, on average, about one-third of the rates in the residential street areas. This occurs in spite of the fact that the housing densities are often much higher on the estates. Remarkably, there are more people in the estates, but the encounter field is much sparser. Figure 13.30 is then a scattergram of the relation between *mean integration* and *mean encounter rate* for each of the six areas. Again, the correlation is very powerful: as integration improves, so the mean encounter rate increases. In other words, just as the predictability of encounter rates from space depends on intelligibility, so the actual level of encounter in an area depends on its degree of integration into the larger-scale system. These two effects between them constitute, we believe, the often observed 'urban desert effect' in new estates, where much space is empty for much of the time and people only appear in space unexpectedly.

One final correlation: the degree of *intelligibility* in an area turns out itself to be highly influenced by the degree of integration. Figure 13.31 is a scattergram showing the correlation between *integration* and *intelligibility* in the six areas. The correlation is weakened, it turns out, by the presence in one of the estates of a main street system line which runs through it and which improves the measured intelligibility of the estate more than it improves its integration. Properly speaking the line should be removed, but if it is, the estate decomposes into two segments. If we remove this estate from the scattergram, however, the correlation again becomes very powerful (Figure 13.32).

Figure 13.29 Concatenated area correlation between intelligibility and real predictability for six systems (r = .5938).

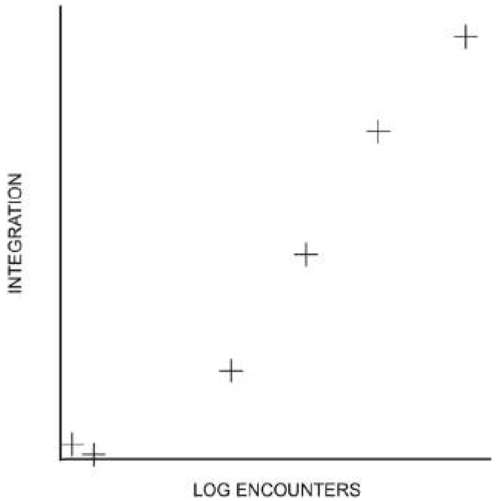


Figure 13.30 Concatenated area correlation of integration and logged encounters (r = .9827).

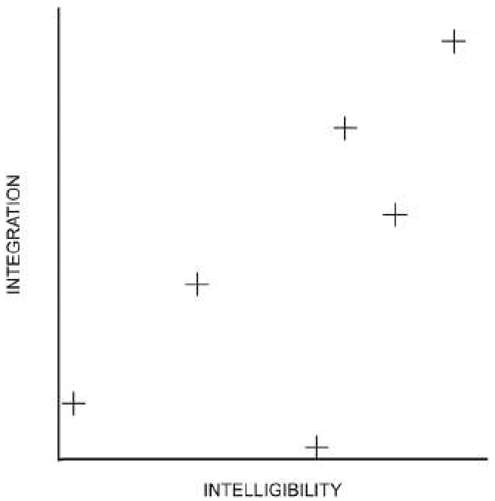


Figure 13.31 Concatenated area correlation between integration and intelligibility for six systems (r = .9401).

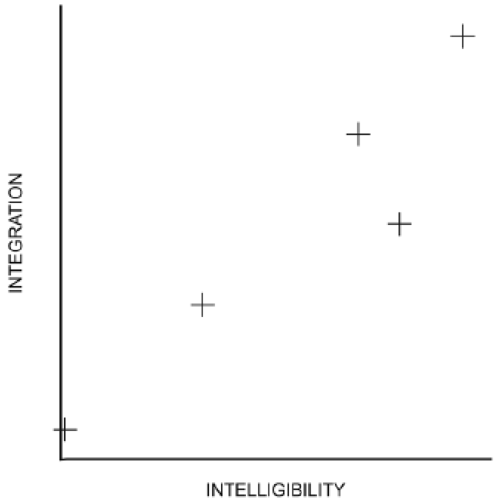


Figure 13.32 Concatenated area correlation between integration and intelligibility for five systems (excluding Packington) (r = .9405).

This interrelationship between integration, intelligibility and predictability has been explored now over a whole range of observational studies and seems to be fundamental. We have also explored the purely spatial interrelationships statistically on a sample of 75 towns and urban areas drawn from many parts of the world to see how far they permit us to use these concepts in building a general theory of the spatial structuring of urban systems, with special reference to the problem of growth. Here again, they show themselves to be very powerful. Taking observational and purely analytic studies together, we are in little doubt that these three properties offer the key to understanding the well-ordering of urban systems, and therefore of the virtual community, which it seems to be their function to construct. The third type of law may therefore turn out to be the most powerful and basic of all the types of law, in the sense that all three types are in some sense a product of this basic propensity for space to construct the field of potential encounter and co-presence we call the virtual community.

The city as socio-spatial artefact

Of course, this is not *all* that urban space does. But, I would suggest, the spatial generation of the virtual community is the precondition for space doing anything at all. What else space does refers, I suggest, to some *transformation* of the virtual community, requiring a parallel transformation of space. To explain this further, I would like to advance two propositions, and then develop them a little on the basis of one more example: the City of London.

The notion of *community* always has within it the theoretically dangerous idea that people are members of a community. This leads the researcher to try to identify the community of which an individual is a member, to discover its spatial limits and to analyse its structural complexities. This overlooks the fact that the *defining feature* of human societies – in contrast to, say, ant or wildebeest societies – is that each individual belongs to *many* types or *transformations* of community. At the very least, every individual will belong to at least one *spatially* defined group – a household group, a village or a university are all instances of such, being defined in terms of the spatial continuity of some domain and the everyday spatial proximity of members – and at least one *transpatially* defined group, such as a clan, a trade, or an academic discipline. These I call 'transpatial' because they link people of a *similar category or kind*, regardless of spatial proximity. Transpatial groups work by *analogy* or *identity* rather than spatial contiguity. They *overcome* spatial separation, and integrate conceptually individuals that are spatially apart.

Now it has become common to argue that the existence and even dominance of transpatial groupings in societies, and their lack of correspondence with spatially defined communities, is evidence for the independence of society from space. Nothing could be further from the case. All transpatial groupings have, at some time or other, a spatial realisation. Any transpatial group that does not from time to time realise itself in space is likely to risk its existence as a group. Thus, clans have ceremonies, trades have congresses and academic disciplines have conferences.

Now it can happen, though rather rarely, that spatial and transpatial groups will coincide with each other, and create a *correspondence* between these two types of human membership. Such systems tend to produce tight, localised and internally hierarchical socio-spatial groups, because both types of membership reinforce each other and tend to render the social group closed. Sadly, the current fashion for the 'territorial' theory of human space in some parts of the Anglophone world, especially the USA, has led to the implicit adoption of the 'correspondence' type as being both basic to human nature (the history of humankind has been rewritten to make this tragic idea appear plausible) and actually desirable as an ideal.

I have dealt with this problem at length elsewhere and it is not my main theme here.[11] What I want to describe here is the more common human situation, where individuals are not only members of many different *transformations of community*, but that these transformations all *co-exist in the same spatial domain*, even though each transformation of community *unfolds in space in a different way*, often according to spatial principles that *contradict* those of another transformation. The city is the archetype of such a spatial domain. Each urban individual lives his life in various kinds of membership, using different spatial principles for each. How these different spatial principles unfold in relation to each other *defines the city* as a *socio-spatial artefact*.

This must sound obscure, so let me work through an example. Figure 13.33 is a map of the City of London dated 1800, before major changes were made to it in the nineteenth century. It is a typical *deformed grid* with a *covering core*, that is, a deformed wheel but without the peripheral spaces. What interests me here is, however, not the spatial structure alone, but the spatial structure in relation to the various public and institutional buildings that form part of it, that is, the buildings that relate to various kinds of community membership.

First, some preliminaries. A striking feature of the City is the extent to which it *lacks* the kinds of spatial and formal ordering which I defined as characterising the cities of *social reproduction*. Thus, the most famous building of all – St. Paul's Cathedral – is axially virtually unrelated to the urban fabric as a whole. The three major buildings at the centre – the Mansion House, the Bank of England and the Royal Exchange – all manage to avoid orthogonal axial lines on their facades in spite of their strong central location, and this is still true today, even though three new major axes are linked to the central area. The major buildings are, it seems, *locationally* but not *spatially* significant in the urban surface. The logic behind this is not hard to find. The City of London is structured in such a way as to generate the dense and continuous encounter field on which its history as a centre of commerce depended, not in order to express the categories of ideological power.

When space does play a role in relation to institutional buildings, however, it is both subtle and unusual. There are two dominant types of institutional building distributed through the urban fabric of the City: guild buildings (that is, the buildings of the old medieval guilds) and churches (most of them rebuilt by

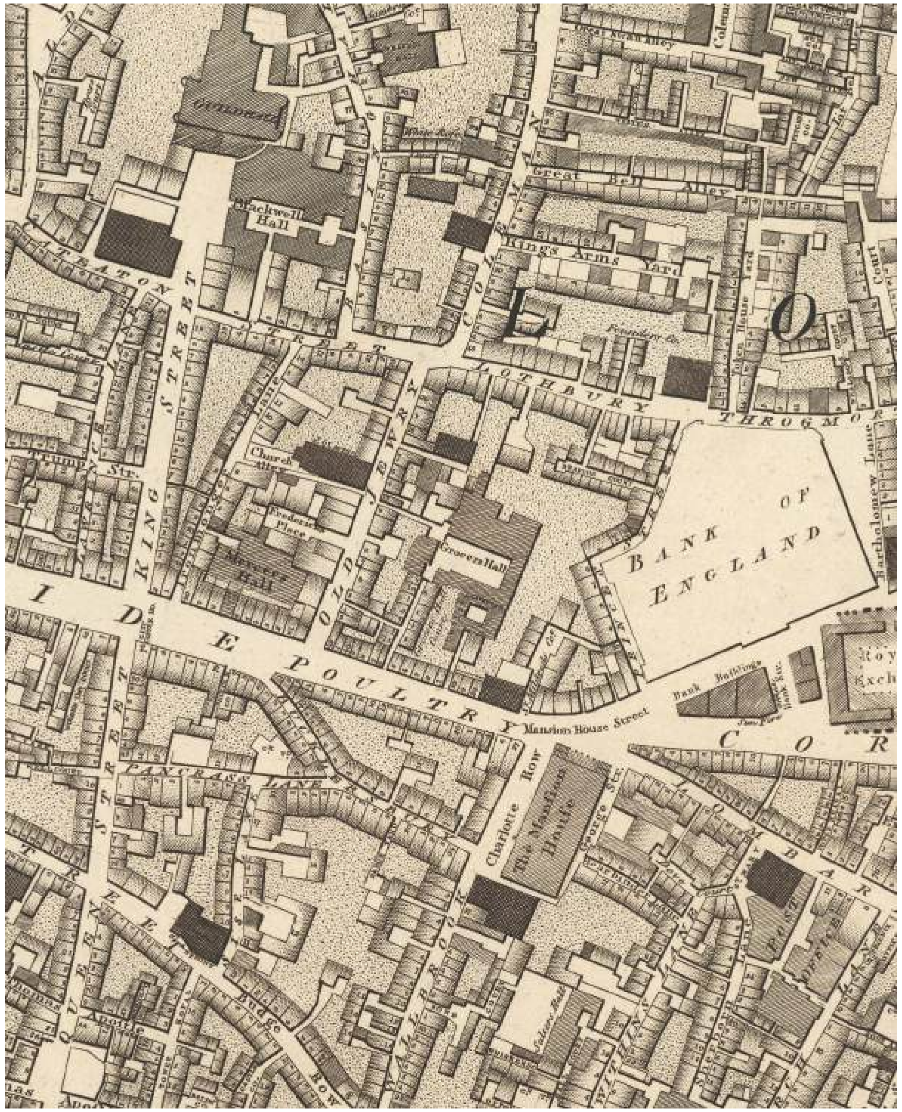


Figure 13.33 Section from Richard Horwood's *Plan of the Cities of London and Westminster*, 1799, showing the location of major guilds and churches within the urban fabric.

Wren after the Great Fire). Each type is spatially distinctive. Churches, unusually – especially considering their architectural importance – are often built into the urban fabric, with other, 'profane' buildings contiguous to them. In spite of this, most are relatively prominent in public space. Many – though by no means all – are located on significant axes, and everywhere their presence in public space is evident, even when partly concealed by neighbouring buildings.

The guild buildings, on the other hand, are much more discreet. One is hardly aware of them from the public spaces. They make little show to the outside, and the courtyards which are a common formal feature tend to be separated from public space by narrow passages. It is only when one arrives in the courtyards that the buildings begin to express themselves as architecture. If one looks further for the halls that constitute the main ceremonial meeting places of the guilds, then this impression is stronger: halls are several steps deep from the outside spatially, even when they have physical adjacency to the courtyard. The guild building seems almost a private realm, and not at all what one would expect of a building concerned to express the identity of a social group in the urban landscape.

Now these two types of building do, of course, embody different concepts of '*social solidarity*' – different *transformations of community* in my terms, with different kinds of membership. The guild building stands for a *transpatial* solidarity. It summons all apothecaries, or fishmongers or drapers, regardless of where they are located in space. Its catchment area is not defined by space, but by *categoric similarities* among individuals. It therefore has some kinship to the sociological concept of '*mechanical solidarity*' as defined by Emil Durkheim.[12] It celebrates the *difference* between an individual and his spatial neighbours by calling him or her, but not *them*, to a separated locale in which this identity can be expressed. Separation is crucial to such identities. Without it, they are in danger of contamination. Only *power* can express social distance without spatial separation – hence the *ideological landscape* of the city of social reproduction.

The churches, in contrast, work on socio-spatial principles which are exactly the contrary. The catchment area of each church is spatial. It belongs to its parish; that is, it belongs to a group that is spatially defined, and takes no account of the categoric differences that might exist among the people of that parish. Insofar as its catchment area extends outside the parish – and it does, of course, do so, since the church is created by, but not restricted to, members of the parish – then it continues to be indiscriminate in this way. In its main interior space – which lacks the careful distancing from outside found in the guild halls, it celebrates not the identity of a categorically defined group, but the community formed by *all who happen to be present*, explicitly *without* regard for the social differentiations that prevail outside. This is the transformation of community that Victor Turner, the anthropologist, has so brilliantly identified as '*communitas*' in his book *The ritual process*, as that form which prevails either in transitional or liminal states in ritual or as temporary states within ritual celebration, and which is characterised by the temporary suspension of all social differentiations and manifestations. Of its essence, 'communitas' destructures society in order to create an undifferentiated, if temporary, community subordinate only to the ritual principle itself.[13]

Typically, the individuals of the City will belong to both types of community, and easily move from one to another, using different socio-spatial principles in each case. But membership does not end there. Both of these memberships are realised *inside* buildings which in some sense represent specific groupings. However, other, less structured and more indeterminate forms of membership are also realised *outside* buildings in the public space of the town.

Again, we may have recourse to Durkheim, this time to his concept of '*organic solidarity*'. Organic solidarity is, according to Durkheim, a form of social cohesion based not on *categoric similarities* (as mechanical solidarity is) but on the *instrumental differences* that arise from the division of labour into specialist trades, and the interdependence that this brings in its train. Such instrumental interdependence depends on a high degree of *spatial integration*, and a high generation of random encounter is needed to support the field of *interaction* required to maintain such a system in operation. Much of this interaction occurs *at the interface between the domain of the individual tradesman and the pattern of public circulation in the streets*. This interface is spatially designed so as to maximise this form of interaction of *individuals with different instrumental identities*. The streets, therefore, insofar as they relate to everyday building entrances, constitute a space of *instrumental interaction*, or of the organic solidarity of individual differences.

In other words, the street interface realises in space the individual differences that form the basis for the group differences celebrated in the guild building. But just as there is an *undifferentiated* transformation of the differentiated *interior* community, realised in Turner's '*communitas*' in the churches, so I believe there is an undifferentiated counterpart to the *interactive* community of the street interface: the transformation of community which results from the *mere non-interactive (or pre-interactive) co-presence of people*. This occurs not so much at the street interface as in the streets themselves, and at the indeterminate interface of an area of streets with the world of strangers, a form of community perpetually growing or shrinking as a function of the capacity of the pattern of space to make this community dense or sparse, structured or unstructured, continuous or sporadic. This transformation of community is of course the *virtual community* itself, perhaps *the* fundamental entity from which the other transformations are possible. We may summarise these four transformations of community in a simple model (Table 13.1).

The essence of this model is that its *spatial* dimensions move strongly *from the local to the global*. They depend on the *non-existence* of local boundaries, and on a *non-correspondence* between the social and the spatial. With the exception of the guild buildings, the boundary of the individual domain is weak, since its aim is to construct an *interface* with the outside. Likewise, the boundary of the *virtual community* is weak, both at the level of the local area and of the City as a whole. At each level, it admits strangers, and thus always tends to growth rather than restriction or stabilisation. Boundaries in the system are associated with the *transpatial* – that is, with the creation of solidarities that do not depend on the structure of the encounter field. The urban space pattern as a whole thus reflects the *globalising* tendency of the *encounter field itself*, rather than an order based on representational buildings. It is geared to the continual *generation* of virtual community, and the *conversion* of the virtual community into the interactive community through the interface of building and street.

Table 13.1 The four transformations of community

Community type	Differentiated	Undifferentiated
Transpatial or inside	*mechanical solidarity* or type differences spatially separated	*communitas* or undifferentiated community bounded and ritualised
Spatial or outside	*organic solidarity* or type differences partially mixed	*virtual community* or undifferentiated community unbounded and profane

Modern transformations

Against this background, we can begin to see that the transformations of the urban landscape by Modernism and its successors have not only a *spatial* logic, but also a *social* logic. We can explore this at two levels: the local urban area; and the global structure of the city. The former is dominated by the idea of *enclosure*; the latter by the urban *monument*. Each claims ancestry in the European city. Yet each *inverts* the common forms of the structure of that city.

Let us begin, as usual, by looking at examples. Figures 13.34a–d are examples taken from an international review of built designs published in 1977.[14] By looking through its pages we can begin to see that certain formal and spatial themes – not those which took pride of place in theoretical debates – have a universality which compels us to regard them as *paradigmatic*, that is, as emanating from a deep, taken-for-granted scheme of assumptions rather than from a specifically articulated theory.

I suggest that we can sum up these paradigmatic ideas in three concepts which between them offer the characteristic modern solution to the fundamental problem of urban design: that of combining *local parts* to form a *global whole*. These three concepts are: *enclosure*, *repetition and hierarchy*. I will discuss these concepts with reference to a pair of examples drawn from the beginning and end of the modern period.

- *Enclosure* is the concept through which the smallest urban *spatial element* is defined. Spatially, it means that a (normally convex) space is defined as distinct from others by being more or less surrounded by buildings. Socially, it means that those buildings and their occupants have a special relationship with that space, a relation that is usually defined in terms of *identity*. Socio-spatially, it means that the inhabitants are expected to identify with each other through the shared identity with the enclosed space. Morphologically, of course, our research on encounter fields shows the argument to be nonsense. Such spaces are invariably empty of people: of strangers, because they are inevitably segregated from the natural movement patterns in the larger-scale urban landscape; of inhabitants, because people forced by space into an over-direct and coercive

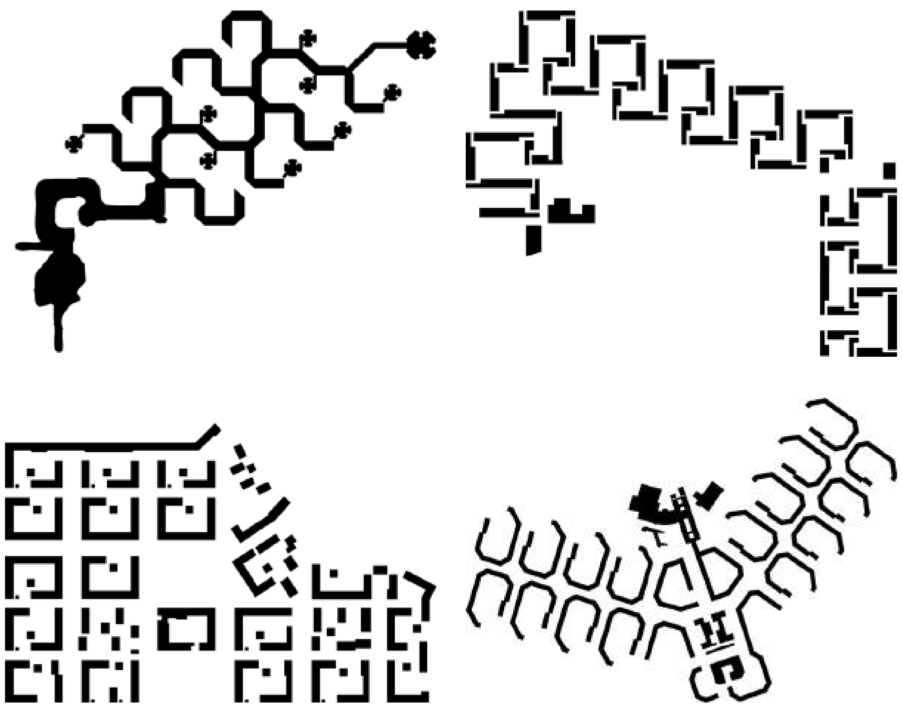


Figure 13.34 An international selection of housing schemes (after Kirschenmann and Muschalek 1980): (top-left) Dimona, Israel; (bottom-left) Bruket, Sandviken, Sweden; (top-right) Steilshoop, Hamburg, Federal Republic of Germany; (bottom-right) Pollards Hill, Surrey, England.

encounter relation in circumstances where the lack of strangers also implies the loss of useful urban anonymity invariably seem to reinforce those behaviours of circumspect avoidance by which we regulate relations with neighbours. The segregated spatial enclosure in effect normally *negates* the spatial community.

- *Repetition* is then the regulating principle by which a larger-scale system is generated out of the local parts. Spatially, it encourages a certain *geometricity* in layouts, and eventually the domination of *spatiality* by geometricity. Socially, the principle is loaded with the implication of a *segmentary* society, that is, a society composed of separate but identical sub-units, whose integration into a larger society is through *symbolic identities* rather than through the *spatial integration* of practical interdependencies. This again resembles the Durkheimian concept of *mechanical solidarity*, with the addition of the idea that the symbolic identities of segmentary groups arise in the first place from a *spatial* identity.
- *Hierarchy* is then the principle by which the aggregation of local parts is given a *globally ordered* form. Spatially, it means that the local principle of enclosure is applied at a higher level. In parallel, the social implication is that the higher-order application of the spatial principle stands for the identification of higher-order social groupings, which, as at the smaller scale, couple this distinct identification with *separation* from the larger world.

The idea of hierarchy as the globalising principle confirms the ascendancy of geometricity over spatiality. It creates an urban surface which has conceptual intelligibility on the drawing board or when seen from the air, but which destroys the spatial intelligibility from local to global which is the essence of the traditional town. It also destroys the virtual community and its potential to convert into the interactive community through the interface of building and open space. Through its over-localisation of the primary spatial element, the enclosure, it limits the scope of the virtual community to a small group of neighbours; and by its removal of the dwelling entrances from the system of truly public space, it limits the interface potential to the same group.

In all respects, then, this modern paradigm of space is founded on a principle of correspondence between groups of people and spaces, and it is this that leads to the dominant values of enclosure and identity as architectural virtues. It is these alleged virtues, however, that most deeply contradict the socio-spatial principles of traditional urbanism. It is this paradigm, not the high-rise building, that has led in our time to the progressive destruction of the urban fabric of our cities through a whole series of architectural movements.

The essence of the modern transformation of the city is the loss of a global order arising from the way local parts are defined. It is this loss of the global that is at the root of our current preoccupation with monumentality. Monumentality is intended to be the means by which the global urban order is reconstituted. But of course without the reconstitution of the urban fabric itself, it can only lead to a caricature, a city in which the monumental structure is added to the city as a separate region in an otherwise spatially fragmented system. Such quasi-urban forms bring to mind the frightening social systems of the proto-urban societies of the ancient Near East and pre-Columbian America rather than anything that can be found within the urban tradition proper.

But even if monumentalism is reintegrated into the urban fabric as a globalising device, it can still only succeed in recreating the city of power and social reproduction, not the city of everyday life and work. Such everyday cities, which are by far the dominant type in urban history, typically do not depend on monumentality for their global form. On the contrary, global form arises from the way in which space is articulated from the local to the global scale. Monumental buildings are then fitted discreetly into an urban fabric whose spatial structure is given by considerations which are at once mundane and marvellous – mundane because they arise from the elementary need for cities to construct an encounter field with a certain structure and density, marvellous because it is this that gives rise to the exhilarating sense of urbanity which pervades everyday life in such cities, and which has given them their dominant cultural character.

The sense of urbanity, I suggest, can only be retrieved in our cities through a new paradigm of socio-spatial organisation, one which emancipates itself from the current obsession with localism coupled to monumentalism and reinstates the principle of global spatial design through the deformed grid, in which local places are differentiated from each other, and yet the whole is intelligible from the parts. Space syntax is, I believe, a necessary part of the means to the understanding and design of such urban systems.

Conclusion

I am aware that I have not offered a full theory of urban space, far less a theory of urban form. I have not dealt with the third dimension, nor with the style and form of individual buildings and have only sketchily dealt with the location of major public buildings in the urban structure. Nor have I dealt with the concept of 'memory', and all that implies for the temporal dimension of cities – though I do believe that the concept of 'intelligibility' opens up some of the configurational dimensions of urban memory. All these are aspects of current space syntax research but each would require a presentation as discursive as this one, and must therefore be kept for another occasion.

However, I must not disguise my belief that, important as these other aspects are, it is the syntax of space itself, especially at the global level, and its relation to the virtual community, that is at the heart of the urban question. This I believe is equally true whether we think of it as a question of understanding and research or as a question of design and action. Space syntax, and its associated theory, is both. It is both research instrument and design tool. In either case, it is a 'thing to think with'.

Notes

1. J. Peponis, C. Ross and M. Rashid, 'The structure of urban space, movement and co-presence: The case of Atlanta', *Geoforum* 28 (1997): 341–358, p. 345.
2. B. Hillier, 'Can architecture cause social malaise?', in *Space is the machine: A configurational theory of architecture* (Cambridge: Cambridge University Press, 1996), pp. 140–141.
3. V. Turner, *The ritual process: Structure and anti-structure* (New York: Cornell University Press, 1969).
4. J. Hanson and B. Hillier, 'The architecture of community: Some new proposals on the social consequences of architectural and planning decisions', *Architecture et Comportement/Architecture and Behaviour* 3 (1987): 251–273, p. 265.
5. R. M. Adams, *The evolution of urban society: Early Mesopotamia and Prehispanic Mexico* (London: Weidenfeld & Nicolson, 1966); P. Wheatley, 'The concept of urbanism', in *Man, settlement and urbanism*, edited by P. J. Ucko, R. Tringham and G. W. Dimbleby (London: Duckworth, 1972), pp. 601–637; H. Carter, *An introduction to urban historical geography* (London: Edward Arnold, 1983).
6. M. M. Webber (ed.), *Explorations into urban structure* (Philadelphia: University of Pennsylvania Press, 1964).
7. J. Rykwert, *The idea of a town: The anthropology of urban form in Rome, Italy and the Ancient World* (Princeton, NJ: Princeton University Press, 1976); Rob Krier, *Urban space*, translated [from the German] by C. Czechowski and G. Black (London: Academy Editions, 1979).
8. A. Rossi, *The architecture of the city*, edited by P. Eisenman and K. Frampton (Cambridge, MA: MIT Press, 1985).
9. Editors' note: the original text also mentions Krier's New Luxembourg as being similar, 'So is Krier's New Luxembourg (Figure 21a)'. We have omitted this image due to it being too unclear for reproduction.
10. B. Hillier, R. Burdett, J. Peponis and A. Penn, 'Creating life: Or, does architecture determine anything?', *Architecture & Behaviour* 3 (1987): 233–250.
11. Hanson and Hillier, 'The architecture of community'; B. Hillier and J. Hanson, *The social logic of space* (Cambridge: Cambridge University Press, 1984).
12. E. Durkheim, *The division of labour in society*, English edn (New York: The Free Press, 1933) (Original edition 1893).
13. V. Turner, *The ritual process* (New York: Cornell University Press, 1969).
14. J. C. Kirschenmann and C. Muschalek, *Residential districts*, translated [from the German] by TST Translations (London: Harper Collins, 1980).

14 Natural movement (1993)

Space and the psychology of natural movement. Introduction to 'Natural movement'

Mahbub Rashid

A fundamental purpose of the built environment is to create, control and distribute movement. Since the built environment is defined by forms, spaces and functions, movement must be affected by them all. Yet, a functional theory generally suggests that movement is affected by functions only without acknowledging the role of form and space in the process.

Published in 1993, the following paper offers a spatial theory of movement that works in tandem with a functional theory of movement. Emphasising the transactional relationships between individuals and their environments, this theory suggests that spatial affordances, such as accessibility and visibility, can help generate movement by setting up differential conditions for interaction and avoidance among individuals of a society. Termed as 'natural movement', such movement is distinct from any movement generated by functions. According to this theory, functions can either increase overall movement by adding movement to natural movement, as we find in many commercial areas of the city; or they can reduce overall movement by keeping natural movement unrealised, as we find in many residential areas of the city.

Using empirical evidence from different environmental settings, this paper successfully makes a case for the theory of natural movement. It uses the techniques of axial map analysis of space syntax to rigorously describe spatial affordance defined by form and space. To remind the reader, an axial line is a straight line of movement and sight in the built environment. An axial map uses axial lines to economically represent the spatial configuration of the built environment. The paper reports statistically significant associations between observed patterns of movement, and accessibility and visibility defined using the techniques of the axial map analysis. These associations generally suggest that movement densities increase in more accessible and visible spaces as defined by space syntax.

The findings reported in this paper can be explained using 'the principle of least effort' and 'the evolutionary principle of prospect and refuge' that scientists often use to explain the transactional relationships between individuals and their environments. The principle of least effort states that, given the option, humans will naturally choose a path of least effort or resistance in their physical as well as mental actions. The principle has been used in widely different fields by many, including linguists, psychologists and behavioural economists.[1] Complementarily, linking spatial affordances to basic psychological needs, the principle of prospect and refuge argues that the probability of survival of pre- and early-humans had increased when spatial affordances provided increased capacity to observe (prospect) while avoiding attention (refuge).[2] Since easily accessible and visible spaces with multiple vantage points would generally support movement following the principle of least effort and of prospect and refuge, it is no wonder that this paper reports that more accessible and visible spaces generate more movement.

The influence of this paper has been widespread. Inspired by the methods and findings reported in it, space syntax researchers have studied different aspects of movement patterns using different modelling techniques. Their studies have helped explain the patterns of different modes of movement – walking, cycling and driving – in relation to spatial affordances. Their studies have also helped explain how the relationships between movement and spatial affordances affect travel time, congestion, crime, diversity, safety, land use, property values, segregation, environmental pollution, health risks and many other aspects of human significance. Additionally, their studies have helped explain the structural logic of space and everyday life in the city, suggesting that the city uses global accessibility to maintain economic efficiency and local accessibility to maintain cultural differentiation.

Besides conducting numerous empirical studies, space syntax researchers have also refined the early techniques used in this paper to improve their abilities to explain movement. For example, more recently they have used the techniques of segment map analysis that replace axial lines with axial segments for a finer description of spatial affordance. Complementing the topological measures used in this paper, they have also used several metric and topo-metric measures of the axial and segment maps, further improving their abilities to explain movement. At the same time, in a recent shift they replaced axial lines with 'natural street' segments. This shift provides computational efficiency at the cost of psychological richness of the axial and segment maps, and takes away the psychological reasoning for natural movement based on spatial affordance.

To conclude, one should note that, since the publication of this 1993 paper, space syntax researchers have primarily explored the first-order effects of space on movement and other related social phenomena (that is, direct relationships). More studies are needed on the second- and third-order effects of space on different social phenomena with movement as a mediating or moderating factor (that is, indirect relationships) for a comprehensive understanding of the social logic and psychology of form and space.

Natural movement

Or, configuration and attraction in urban pedestrian movement

Bill Hillier, Alan Penn, Julienne Hanson, Tadeusz Grajewski and Jianming Xu

Originally published: Hillier, B., Penn, A., Hanson, J., Grajewski, T. and Xu, J. 1993. 'Natural movement: Or, configuration and attraction in urban pedestrian movement'. *Environment and Planning B: Planning and Design*, 20, 29–66.[3]

Existing theories relating patterns of pedestrian and vehicular movement to urban form, characterise the problem in terms of flows to and from 'attractor' land uses. This paper contains evidence in support of a new 'configurational' paradigm in which a primary property of the form of the urban grid is to privilege certain spaces over others for through-movement. In this way, it is suggested that the configuration of the urban grid itself is the main generator of patterns of movement. Retail land uses are then located to take advantage of the opportunities offered by the passing trade and may well act as multipliers on the basic pattern of 'natural movement' generated by the grid configuration. The configurational correlates of movement patterns are found to be measures of global properties of the grid with the 'space syntax' measure of 'integration' consistently found to be the most important. This has clear implications for urban design, suggesting that if we wish to design for well-used urban space, then it is not the local properties of a space that are important, in the main, but its configurational relations to the larger urban system.

Introduction: Attraction and configuration

Quantitative methods for predicting pedestrian movement in urban space have conventionally been adaptations of the models employed in vehicular studies. The trip-generation potential of built forms (buildings or urban blocks) is seen as the key quantity, congestion as the most likely problem and the scaling of local pedestrian space to match attraction as the main design aim.[4] We might call this the *attraction* theory of pedestrian movement: movement is seen as being *to and from* built forms, with differing degrees of attraction, and design is seen as coping with the local consequences of that attraction.

It follows that attraction theories say little about the spatial *configuration* of the urban grid, that is, about the way in which the spatial elements through which people move – streets, squares, alleys and so on – are linked together to form some kind of global pattern. But it is easy to show that, theoretically at least, configuration can have effects on movement which are independent of attractors. For example, in the simple layout shown in Figure 14.1a, all journeys from 'side street' origins to 'side street' destinations must pass through one or more segments of the 'main street', giving a movement pattern in which the more central segments of the 'main street' are likely to be the best used, and the peripheral segments the least. This would remain the case, regardless of any metric deformation of the pattern, provided the topology was retained. In the slightly more complicated case, shown in Figure 14.1b, there is a relation between configuration and movement which is less deterministic, in that assumptions about metrically or topologically shortest routes are required. The case is also more complex in that spaces other than the 'main street' are involved. For example, the two most central vertical elements, one above and one below the 'main street', would be on more shortest routes than more peripheral, vertical elements.

These effects are of configuration on *through-movement*, and are seen, if we consider the layout as a system of possible routes. But, if we consider the layout as a system of origins and destinations, it becomes clear that configuration may also be implicated in *to-movement*. For example, in Figure 14.2a the 'central square' offers a metrically, or topologically, more accessible destination than the other spaces in the layout. This also holds for the more central elements in the improbable layout of Figure 14.2b. To the extent that accessibility of destinations is a factor in the choice of destinations we could, again, expect to find effects of configuration on movement.

On the face of it, then, configuration may have effects on both through-movement and to-movement in urban grids, which are independent of built-

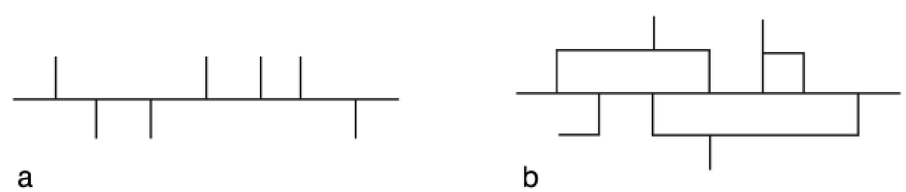


Figure 14.1 (a) The more central segments of the 'main street' are likely to be the best used, and the peripheral segments the least. (b) The two most central vertical elements, one above and one below the 'main street', would be on more shortest routes than more peripheral vertical elements.

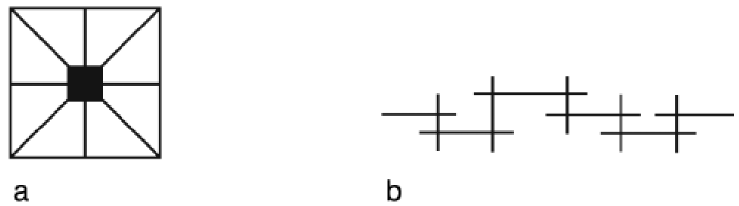


Figure 14.2 The ‘central square’ (a) and central elements (b) offer a metrically or topologically more accessible destination than the other spaces in the layout.

form attractors and perhaps, to some extent, also of metric properties. Are such configurational priorities found in real urban grids? And if so, do they matter? Common sense and common practice would surely insist they are and they do. Urban grids are almost invariably conceptualised as some kind of ‘spatial hierarchy’, in which different kinds of configurational priority are assumed to be associated with different degrees of functional importance. The notion of some kind of configurational structure with functional implications is, it seems, usually present in our notions of urban form, even though its theoretical and formal articulation is unsophisticated.

Let us follow common sense for a moment and suppose that configurational priorities are both present in urban grids and important enough to have significant effects on movement patterns. What would follow from this? It would surely follow that, as an urban system evolved, the distribution of built-form attractors might itself be influenced by these priorities. For example, spaces which the grid configuration prioritised for through-movement might, for that reason, already have been selected as good locations for ‘passing trade’ land uses. Other types of land use might equally have sought to minimise the possible interference of through-movement. Similarly, topologically or metrically accessible locations may have been preselected for types of land use where this was a useful asset, and vice versa.

If, then, at any stage of the evolution of the urban system we were to investigate movement patterns and found agreement between movement rates and the presence of attractors, it would clearly be unwise to assume that movement could be explained by attractors until we were sure that the configurational properties had not influenced both the presence of movement and the presence of attractors. If we then found that configurational properties were in agreement, both with movement rates and with attractors, how should we proceed? We would seem to have the familiar problem of needing to distinguish the respective ‘causal’ effects of two variables which are both correlated with a third, and which are also correlated with each other.

The matter may not, however, be as difficult as it looks. In a situation where movement, configuration and attraction were all in agreement, there would

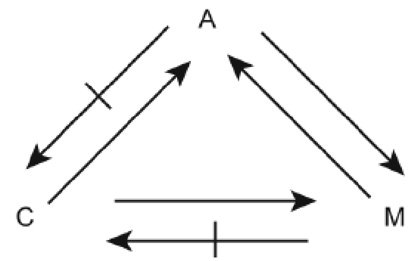


Figure 14.3 A is attraction, C is configuration, M is movement. Attractors and movement may influence each other, but the other two relations are asymmetric. Configuration may influence the location of attractors, but the location of attractors cannot influence configuration. Likewise, configuration may influence movement, but movement cannot influence configuration. If strong correlations are found between movement and both configuration and attractors, the only logically possible lines of influence are from configuration to both movement and attractors, with the latter two influencing each other.

be powerful, logical reasons for preferring configuration as the primary ‘cause’ of movement. Logically, the presence of attractors can influence the presence of people, but it cannot influence the fixed, configurational parameters which describe its spatial location. Similarly, configuration may affect movement, but configurational parameters cannot be affected by it. If we find a strong degree of agreement among all three, therefore, then it must either be the result of chance or the result of configuration having influenced both the pattern of movement and the distribution of attractors. In other words, if we find all three in agreement, then we are compelled to assign causal primacy to configuration. We may clarify this through Figure 14.3.

In real urban situations, we may also find that the issue of attraction versus configuration is less of an issue. For example, in many urban areas, attractors tend to be clustered in specific locations – shops in a high street, for example – and, given our premises, we might expect the selection of these locations to have been configurationally influenced. In such cases, we would expect the grouped attractors to act as multipliers on a basic movement pattern generated by the configuration. Alternatively, we might find urban areas where the effects of attractors were equalised by their being evenly distributed throughout the system, for example, in a residential area. In such cases, it would be reasonable to expect that configuration would be a more obviously dominant influence on movement. It would only be where the pattern of attractors was both strong, unequal and distributed, without regard for the configurational logic of the urban system, that we would expect to find a lack of influence of configuration on movement.

The theory of natural movement

In this paper, it is proposed that these theoretical premises describe the real state of affairs. In urban systems, configuration is the primary generator of pedestrian movement patterns and, in general, attractors are either equalisable or work as multipliers on the basic pattern established by configuration. This is not to say that in all situations the greater proportion of movement is generated by configuration. On the contrary, it will often be the case that the multiplier effect of attractors far exceeds the effects of configurations. The argument is that configuration is the primary generator and, without understanding it, we cannot understand either urban pedestrian movement or the distribution of attractors or, indeed, the morphology of the urban grid itself.

Because movement generated by the grid configuration is so basic, we suggest it should be identified by a special term. We propose the term *natural movement*. Natural movement in a grid is the proportion of urban pedestrian movement determined by the grid configuration itself. Natural movement, although not always quantitatively the largest component of movement in urban spaces, is so much the most pervasive type of movement in urban areas that without it, most spaces will be empty for most of the time. It is also the most consistent, so much so that it is difficult to avoid the inference that natural movement is the raison d'être of the urban grid itself. Urban grids seem to be structured in order to create, by the generation and channelling of movement, a kind of probabilistic field of potential encounter and avoidance.

This is not to say that natural movement is not a culturally variable phenomenon. On the contrary, it takes different forms in different cultures, reflecting the different spatial logics of the urban grid. Urban grids are cultural products because they create, through natural movement, encounter fields with different structures. These differences are primarily composed of different degrees and types of probabilistic interface between different categories of person: inhabitants and strangers, men and women, adults and children, social classes, and so on.

What is invariant about natural movement is the logic that links spatial configuration to movement. The key element in this relation is that natural movement is a *global* property of a configuration in that it responds to configurational parameters which relate each spatial element to every other element in a system, which may be several kilometres in diameter. Natural movement is only secondarily influenced by local spatial properties, such as those which describe the relation of each space to its neighbours, or the neighbours of its neighbours. Where design is over-localised, as has often been the case in twentieth-century urban design, then the natural movement pattern will be disrupted, and space will tend to become radically underused.[5]

If the theory of natural movement is right, then it would follow that the modern tendency to see the urban grid as a by-product of other processes, even as an epiphenomenon, is misconceived. The grid is itself implicated in the generation of urbanism and its functional logic. Natural movement shows that movement is fundamentally a morphological issue in urbanism, a functional product of the intrinsic nature of the grid, not a specialised aspect of it. As such, the question of movement, and of space use in general, cannot be separated from the question of urban form itself.

Space syntax and natural movement

Natural movement has come to light as a formal and empirical phenomenon through the application of new techniques of configurational analysis known as 'space syntax' to the analysis of the local and global structure of the urban grid, and their coupling to simple techniques for observing space use and movement. It is noteworthy that the space syntax techniques, which post-dict[6] natural movement, were not originally aimed at modelling movement but at understanding the morphological logic of urban grids, especially their growth. The theory of 'natural movement' is literally a by-product of a research programme with different aims. It was only the discovery of the pervasive relation between configuration and movement that alerted us to the possibility that movement might be as fundamental to the morphology of urban grids as we now believe it to be.[7]

Our aim in this paper is to set out the evidence for natural movement and to argue that, even on present evidence, configurational models ought to be brought into play, alongside attraction models, in the analysis and design of urban systems. The paper is in four parts. First, the concepts and techniques of space syntax are introduced, with reference to other texts where they are more fully discussed, where necessary. Second, a selection of case studies is presented showing the relation between configurational parameters and movement. Third, some examples of probabilistic interfaces are sketched. Fourth, a number of inferences are drawn for design, and for the morphological theory of urban form, with special reference to the question of scale.

The deformed grid

The grid of a town or city may be defined as the system of space of public access created by the way in which buildings are aggregated and aligned. For clarity, we may represent an urban grid by reversing usual conventions and showing the space as black and built 'islands' as white (Figure 14.4a). This has the useful effect of bringing the grid to the foreground as the prime object of analysis.

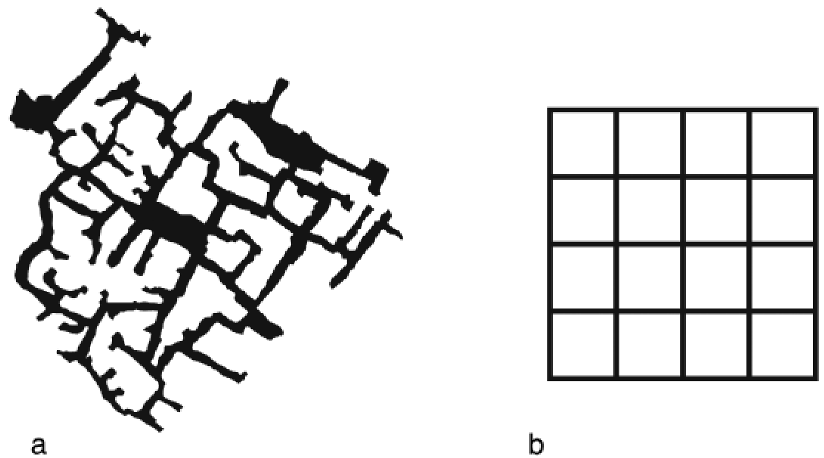


Figure 14.4 (a) A typical deformed grid town is characterised by apparent irregularity. (b) An ideal grid is characterised by geometric regularity.

Most urban grids are, and always have been, *deformed* grids, characterised by apparent irregularity (Figure 14.4a), rather than *ideal* grids, characterised by geometric regularity (Figure 14.4b). This presents obvious difficulties for analysis, because most of our special concepts are geometric. Urban grids are also, by definition, *continuous*. This creates a second difficulty. How is the urban grid to be represented as a set of discrete elements to make configurational analysis possible?

The obvious answer is to use the customary topological representation of the grid as a graph in which route intersections are the nodes and route segments the edges. There are two problems with this representation. First, it makes grids look too similar to each other. It is, for example, a simple matter to design a highly irregular arrangement of 16 urban blocks, but with a node graph isomorphic to that of Figure 14.4b. Second, few interesting consequences, theoretical or empirical, seem to follow from this representation.[8] What seems to be needed is a representation, less general than the topology of the node graph, yet less precise than the geometry of the form itself.

However, we may usefully approach deformity and continuity in grids by making comparisons to geometric regularity. The ideal, orthogonal grid of Figure 14.4b is one in which islands of built forms, with identical shape and perfect alignment, create a grid structure in which all lines of sight and access cross the grid, and pass everywhere through space of uniform width. Spatially, we might say that in the ideal grid, one-dimensional, or axial, elements are as extended as they can be, and so are two-dimensional, or convex elements. Every point in space belongs both to lines and convex elements which cross the grid unimpeded from side to side.

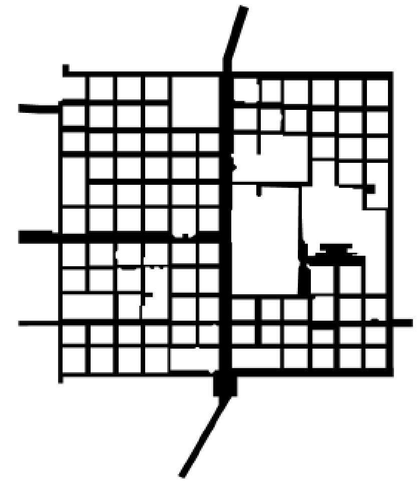


Figure 14.5 Timgad, a Roman grid town in which overall geometric regularity is maintained but spaces are varied in width and alignments are broken by making some islands vacant, or partly vacant, and others are used to block lines of sight and access.

The deformed grid is different. First, in that the shaping and alignment of islands breaks the continuity of lines of sight and access across the grid; and, second, in that spaces vary in width as one passes along lines. More formally, we might say that two types of deformity have been introduced: one-dimensional, or axial, deformity by the breaking up of lines of sight and access; and two-dimensional, or convex, deformity by variations in the width of spaces.

In the example shown in Figure 14.4a, deformity is clearly associated with the lack of geometric order. But this is not necessarily the case. For example, in the Roman grid shown in Figure 14.5, overall geometric regularity is maintained, but spaces are varied in width, and alignments are broken by making some islands vacant, or partly vacant, and using others to block lines of sight and access. Deformity in both types of grid can be described by comparing the degree of axial and convex break-up with a perfectly regular grid.

Axial graphs, node graphs and axial maps

The fundamental proposition of 'natural movement' is that movement in an urban grid is determined, all other things being equal, by the distribution of a configurational quantity called 'integration' in the axial graph, of the axial map, of that grid.[9] The 'axial map' of an urban grid consists of the longest and fewest straight lines that can be drawn through the spaces of the grid so that the grid is covered. 'Covered' means that all rings of circulation are completed and all convex elements passed through. The axial graph is the graph in which the lines of the axial map are the nodes and the intersections of the lines are the edges.

The axial graph is thus a representation of the grid in which the ordering of nodes is lost. All information in the graph is of line-to-line relations, irrespective of the ordering of nodes. The axial map, however, does have information on the ordering of nodes. The topology – though not of course the geometry – of the axial map can thus be reconstructed by adding the node graph to the axial graph. Practically speaking, the axial graph is sufficient for the post-diction of movement, although this is not to say that a more refined form of analysis, in which node-to-node information was also included, would not yield even better results.

The 'axial map' can be constructed automatically, by computer, using a series of rule-based algorithms given the geometric description of the building block forms. It is also practicable to draw the axial map manually by the 'eyeball method'. The map is transformed into a matrix representation, where each axial line is numbered, and an incidence matrix is compiled of connectivities between lines – a line is said to be connected when it crosses or intersects another.

In this way, a geometric, building block plan is transformed into a matrix, or graph representation, where each node of the graph is an axial line (a direct line of sight and access) and each link describes the other lines that are visible and accessible from it. Essentially, the basic description already describes something that is perceptually important, and relates to how one might understand and move around a configuration.

There are a number of measures of the graph which can now be used to describe configurational properties of the grid. The simplest are those that describe the local properties of a node in the graph (that is, an axial line): 'connectivity', for example, merely measures how many other nodes are directly accessible from it (in graph theory this is called the 'valency' or 'order' of a vertex); 'control value'[10] measures the degree to which a node 'controls' access to and from its neighbours. It is calculated by summing the reciprocals of connectivities between neighbours as follows: if the line is the only connection to a neighbour, it acquires a value of 1 from that neighbour; if it is one of two connections, then it acquires ½; if one out of three, then ⅓; and so on. In effect, therefore, control value indexes the amount of choice each line represents for each of its neighbours.

A similar measure, taking into account the relations between each space and the whole system, is 'global choice', which indexes how often each line is used on topologically shortest paths from all lines to all other lines in the system. It thus also indexes how often each line is visited, on random journey simulations, through topologically shortest paths in the system. But empirically, by far the most important global measure is called 'integration', which measures the mean depth of every other line in the system from each line, in turn, relativised with respect to how deep they could possibly be with that number of lines, then standardised.[11] The most integrated lines are those from which all others are shallowest on average, and the most segregated are those from which they are deepest.

A key property of interest is how the various configurational variables are distributed in the urban grid. This can be shown graphically by drawing 'core maps' of, for example, the 10 per cent most integrated lines and the 50 per cent least integrated (most 'segregated') in a system. In most towns, or urban areas, integration core maps will pick out the main thoroughfares and shopping areas, whereas the least integrating will tend to pick out areas with primarily residential functions, that is integration cores seem to offer a graphic realisation of some morphological 'deep structure' of a town or urban area (for example, see Figure 14.9a below). The reason why this may be so will become clear in the case studies of the relationship between spatial configuration (especially the distribution of 'integration') and movement which now follow.

Case study 1: King's Cross

The first case study will also introduce the method through which configurational properties and movement patterns are examined conjointly. The study is of the King's Cross development site, and was carried out on behalf of the Railway Lands Community Group, with funding from Foster Associates, masterplanners for the site, and the developers, the London Regeneration Consortium. The aim was to analyse the urban configuration of the area around the King's Cross site and how it related to existing patterns of movement and space use, then to advise on the implications of this for the agreed design objective of knitting the new development into the surrounding urban fabric of Camden and Islington. Only the analytic aspects of this study are reviewed here.

The spatial structure of the area

Figure 14.6 is a map of the area surrounding the King's Cross site, showing the modelling area bounded by Tufnell Park in the north, Upper Street in the east, Regent's Park in the west and Holborn in the south, covering an area of around 12 km^2. Figure 14.7 shows the street network of the area in black, including the internal, ground-level spaces of several housing estates in the vicinity of the site. The housing estates are easily picked out by the dramatic reduction of spatial scale and increase in spatial complexity. Figure 14.8 is an 'axial map' of Figure 14.7, as digitised into the computer. Syntactic analysis was carried out at two levels and in two modes: of the whole area (the 'large area'), and of a smaller area bounded by Camden Road (north), Caledonian Road (east), Euston Road (south) and Eversholt Street (west). The idea of varying the scale of analysis is to check for any analytic effect that might result from the choice of boundary.

The scale of the smaller area is based on a rough estimate of the normal pedestrian catchment area for the site, and the larger area on a rough estimate of 'catchment area of the catchment area'. Previous studies have invariably shown that

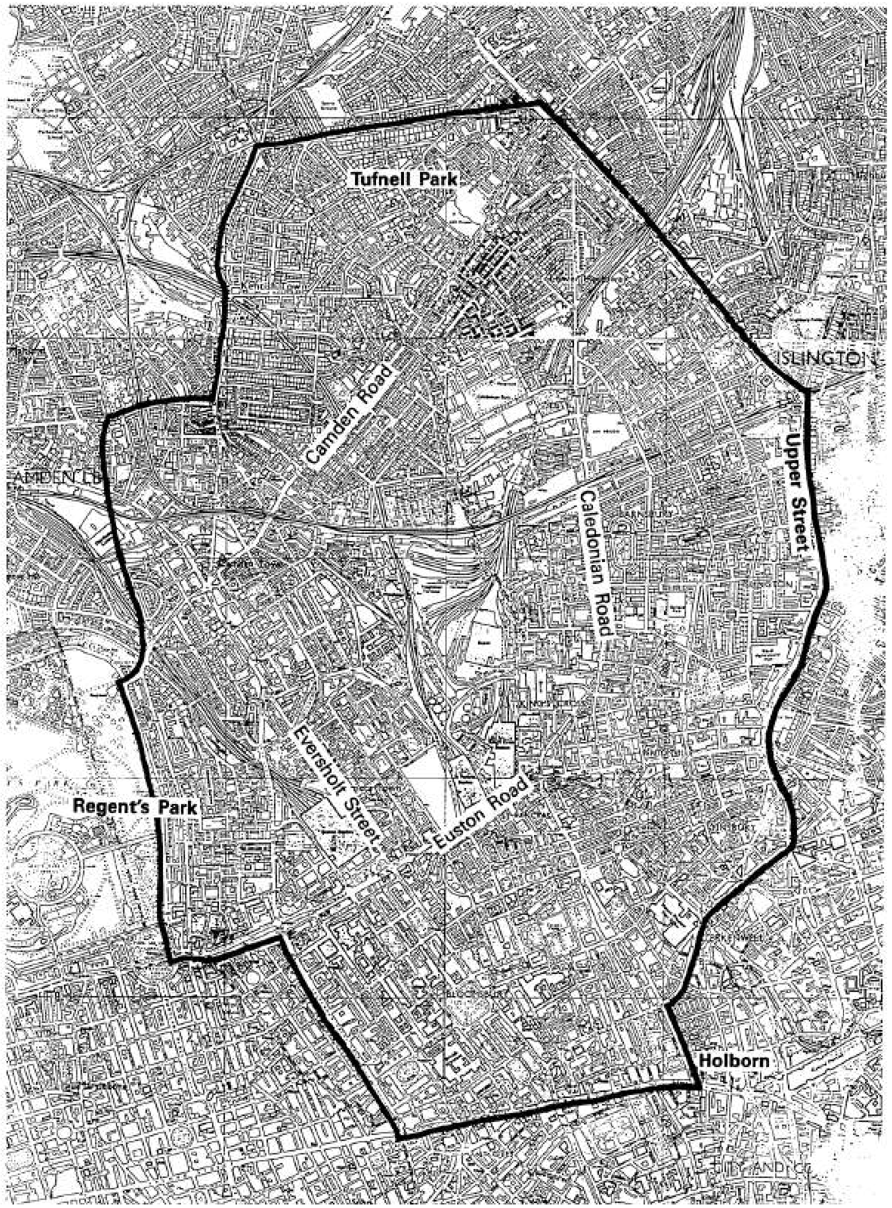


Figure 14.6 Map showing the King's Cross site and the surrounding area, circa 1979.

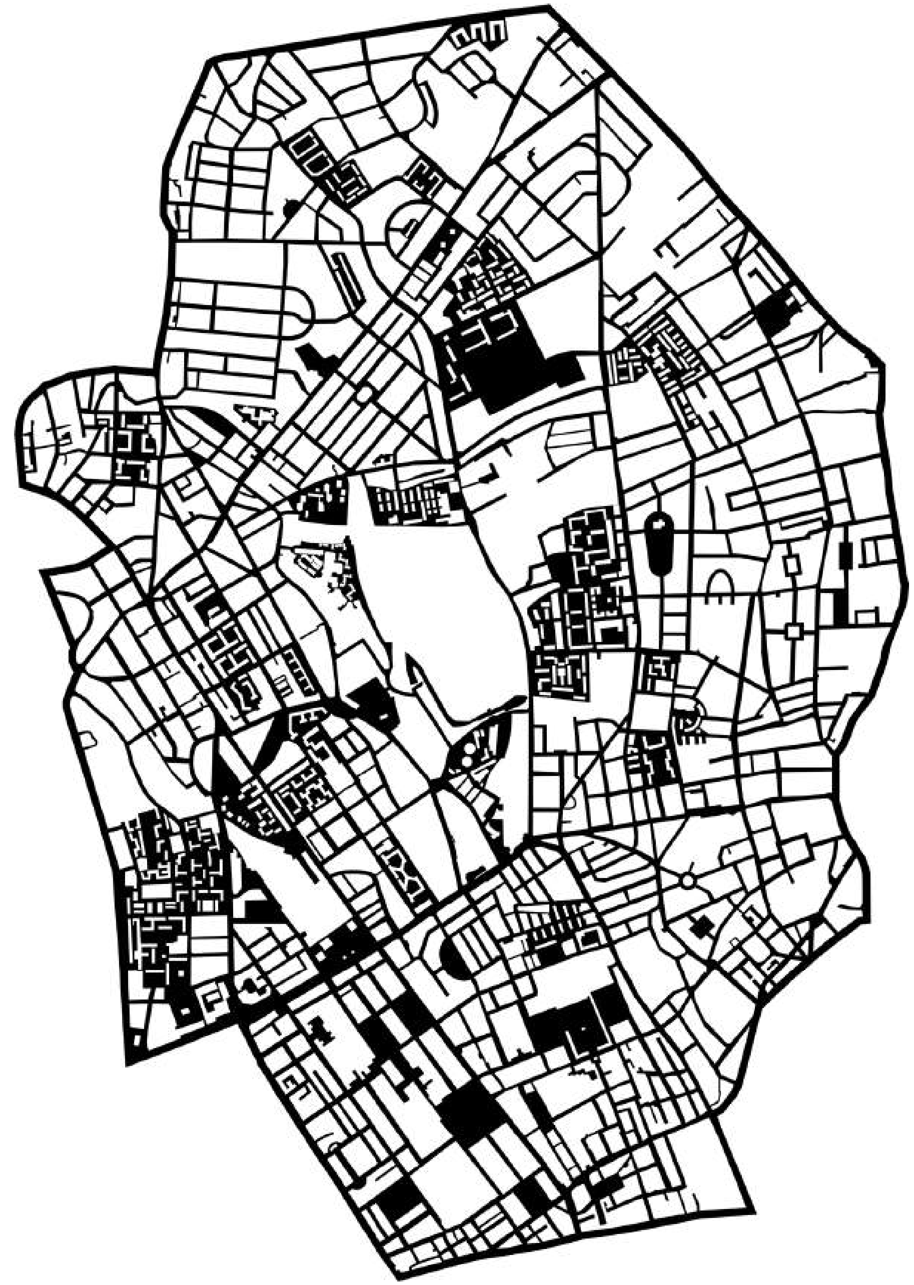


Figure 14.7 The street and open space pattern in the larger King's Cross area, including housing estates in the vicinity of the site.

Figure 14.8 Axial map of the larger King's Cross area, including housing estates in the vicinity of the site.

movement patterns are globally, not locally, determined and that post-diction is only possible from syntactic parameters by setting the area into a reasonably large urban context. It follows that, if we are to predict natural movement within a development site, we must first post-dict movement in the catchment area, and, if we are to post-dict movement patterns within the catchment area, then we must know how the whole of the catchment relates to its own catchment area. A further benefit of the 'catchment area of the catchment area' approach is that it also displaces any 'edge effect' in the spatial analysis into the outer reaches of the larger system, away from the area of prime interest in the more immediate vicinity of the site.

Syntactic analysis of the axial map produces two kinds of output: alphanumeric data in the form of line numbers with spatial parameters assigned to each; and graphic data in the form of 'core' maps in which lines are coloured up in accordance with their value on the various parameters. The latter are, of course, crude approximations compared with the former, but they do allow the designer an immediate, intuitive grasp of the structure of the spatial pattern of the area.

From the point of view of natural movement, the key result of computation is the distribution of integration in the axial map. Graphically, the most informative strategy is to look at an axial map in which the integration values of lines are represented by the colours of the spectrum from red, for the most integrated, through to indigo, for the most segregated. This allows one to form a picture of how integration spreads through the urban structure. Figure 14.9a shows the distribution of integration in the largest system (though in this case in black and white with the most integrated lines in heavy black through to the most segregated in light grey). Coupling the graphic representation with the alphanumeric output, we can trace the actual *order* in which lines integrate the system, and the degree to which each does so.

It is best if we first consider the area without the interiors of its housing estates, because this will give the truest picture of the fully public space of the area as it has evolved. If we do so, the most integrating line in the area is the section of Euston Road between Tottenham Court Road and King's Cross. The second is the line covering Eversholt Street and the upper part of Southampton Row, and the third is the southernmost section of York Way, up to Copenhagen Street. From there the core moves to Camden High Street, Pentonville Road, Royal College Street and Tottenham Court Road. Five of the top seven lines thus run north-south, with only Euston Road and Pentonville Road running east-west.

The core of the area around the King's Cross site thus has not only a strong north-south bias, but also a bias towards the south and west. Two of the next three lines, Crowndale Road and Copenhagen Street, do run east-west, but then the north-south bias takes over again, as does the bias to the south and west. The Barnsbury section of the Caledonian Road and Camden Road are relatively weak for such long lines, coming in 15th and 18th positions respectively, which emphasises the comparative weakness of the core to the east and north of the site.

Figure 14.9a Pattern of integration in the larger King's Cross area excluding housing estates, with the most integrated lines in heavy black through to the most segregated in light grey.

Figure 14.9a also shows that the distribution of integration to the west and south of the site has a different character from that found to the east and north. To the east and north, the core is linear: a small number of linearly connected and branching street sections forming major routes, but without lateral development linking these routes into well-structured subareas. To the west and south, the core is more grid-like, defining local area, core grids for Camden Town, Somers Town and the Argyle Square area.

We can thus say that, to the east and north of the site, the core is sparse and linear, and does not construct subareas, whereas, to the west and south, the core is denser and more grid like, and creates identifiable subareas. More simply, the west and south have a property which we might call *grid-integration*, which is a common characteristic of urban centres and subcentres, whereas the east and north have the more restricted *line-integration*, which is often found in areas with a less developed urban character.

If we now add in the housing estates, Figure 14.9b, we find a further structural phenomenon: certain estates, especially the two immediately north and northwest of the King's Cross site, form solid blocks of segregated lines, unrelieved by either integrating or intermediate lines. Typically, street-based urban subareas will mix and intersect integrated, segregated and intermediate lines. The 'structural segregation' which often characterises modern housing estates appears to be very much a twentieth-century phenomenon.

Observing movement

The next stage of the study was an extensive first-hand investigation of space use and movement. In space syntax studies, an observation technique is used in which observers walk at about 3.75 mph (5.5 kph) along selected routes of about 20 or 30 line segments (mixing integration and segregation as far as possible) and count the people they pass who are moving, or static, on the same line, that is people crossing the line rather than walking along (who are discounted). Distinctions are made between men and women, and children, the last being estimated as being 16 years of age or less. Children who are not moving independently are not counted. Routes are observed between 20 and 30 times, taking care to cover all times of the day by gathering approximately equal numbers of observations in each of five standard time periods: 8–10am, 10am–12pm, 12–2pm, 2–4pm, and 4–6pm. In this case, the extreme time pressure on the study meant that observations were conducted in an extraordinary variety of weathers, from hot to cold and from fine weather to fairly heavy rain. Experience shows – and this study confirms – that weather has relatively little effect on natural movement, although it does affect static behaviour considerably.

Routes were selected in 10 areas around the site, including three housing estates, and covering a total of 239 street sections. The routes are shown as the dark

Figure 14.9b Pattern of integration in the larger King's Cross area including housing estates, with the most integrated lines in heavy black through to the most segregated in light grey.

lines in Figure 14.10. The figure alongside each line is the adult movement rate for that line, that is the average number of adults passed on that line for all observation periods, standardised for the length of the line and observed to a norm of the number of people per 100 m walked. This also approximates to the rate-per-minute of walking time. It is useful to set these figures against the urban norm for residential streets of about 2.7 persons per hundred metres per minute (phm/min).

Positive and negative attractors?

Mean movement rates for each area are tabulated in Table 14.1, with and without main shopping streets. Even without shopping streets, there are very large differences in these rates in the different areas, and there seems to be a pattern to these differences. Of the seven urban areas, which are not estates, those to the east and north of the canal (areas 2 – East, 3 – Northeast, and 4 – North) have an overall mean encounter rate of about 1.64 adults phm/min, whereas those to the west and south (areas 5 – West, 6 – Southwest, 7 – South, and 1 – Southeast) have a mean of 5.51 phm/min. These differences seem to correspond broadly to the difference between 'grid-integration' areas and 'line-integration' areas. They also correspond broadly to the pattern of attractors. The highest average rates are found in the South area, immediately below the entrances to King's Cross and St

Table 14.1 Mean encounter rates (in people per 100 m per min) for King's Cross area and subareas

	Moving	Static	Children
Whole street system	4.12	1.27	0.20
South area	7.27	1.47	0.22
+ Euston Road	9.02	2.16	0.21
Southeast area	2.71	0.86	0.19
+ Caledonian Road	4.22	1.01	0.20
Southwest area	2.68	1.02	0.20
+ Euston Road	4.80	1.73	0.19
West area	3.25	0.95	0.18
+ main streets	5.47	1.87	0.28
East area	1.58	1.10	0.30
Northeast area	1.81	0.89	0.22
North area	1.58	0.27	0.23
Maiden Lane estate	0.27	0.10	0.88
Elm Village estate	0.28	0.22	0.73
Bemerton estate	0.57	0.50	0.44

Figure 14.10 Observed spaces around the King's Cross area inscribed with the numbers of moving adults per 100 m.

Pancras railway stations, and the second highest in the West area, which includes the major shopping centre of Camden High Street.

If we look in more detail at the movement rates set out in Figure 14.10, it certainly seems to be the influence of attractors on the pattern that is initially striking. Not only are the highest levels of movement in the immediate vicinity of station entrances and Camden High Street but, more strikingly, movement levels appear to fall away in all directions as lines recede from these major attractors. For example, in all cases where two sections of the same street have been observed in the vicinity of the King's Cross station entrance or Camden High Street, the segment more distant from the attractor has substantially lower movement rates than the closer one. This effect is so consistent that it is tempting to define strong attractors in terms of this falloff, as much as in terms of unusual concentrations of movement.

If we do so, however, then it seems we must also acknowledge the existence of two kinds of *negative* attractors, corresponding to the two aspects of this definition. First, all three housing estates appear as negative attractors in that their average movement rates are, at 0.365 phm/min, about an order of magnitude lower than the general urban average for residential streets of 2.7 phm/min, or the King's Cross area, overall average of 4.12 phm/min. The estates are as conspicuously empty of movement as the attractor streets are full.

Second, the King's Cross site north of the stations, at present, constitutes a large hole in the urban fabric, and it is striking that movement rates all around the site, north of the stations, are very low, and seem to become lower as the periphery of the site is approached. Even in the highly integrating area of the York Way–Copenhagen Street intersection, rates are well below the urban average, although not as low as the other, more segregated, sides of the site. In other words, the second attractor property, that of a falloff in movement rates receding from the attractor, seems to be inverted for the King's Cross site: movement rates fall off as the 'negative attractor' is approached. We will return to these phenomena later to discuss how far they should be seen as attraction effects or configuration effects, after we have discussed the relations between configurational parameters and movement rates.

Spatial configuration and movement

Figure 14.11a is a scattergram plotting integration values against movement rates for all 239 observed spaces, with integration values read from the largest system. Figure 14.11b is a plot of the same integration values against the natural logarithm of movement rates. Figure 14.12 consists of plot connectivity, control value and global choice, against the logarithm of movement rates.

Two points are particularly worth noting about these scattergrams and correlation coefficients: first, the degree to which integration outperforms other variables; and, second, the degree to which taking the logarithm of movement rates linearises the relationship with integration. Both are consistently found in space syntax studies of urban movement: movement is usually post-dicted best by integration against logged movement rates.

In view of the strength of attractor effects already noted, the strength and quality of the scattergram seems to be a quite remarkable result, all the more so in view of the fact that, for most lines, only one or two segments were observed, and observations were carried out at all times of day and through extremes of weather conditions. Multiple regression confirms the superiority of integration, and shows that adding the effects of other variables does little to improve the overall prediction. In spite of the presence of attractors, it seems, the relationship between

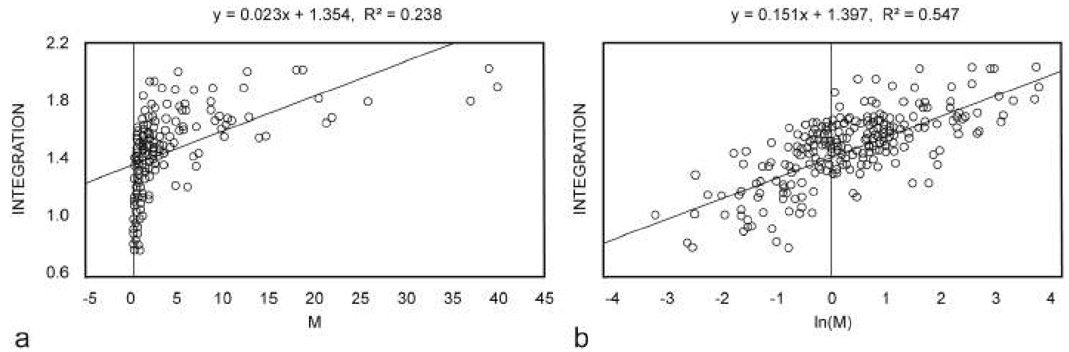


Figure 14.11 (a) A scattergram plotting integration values against movement rates (M) for all 239 observed spaces, with integration values read from the largest system. (b) The same integration values plotted against the natural logarithm of movement rates.

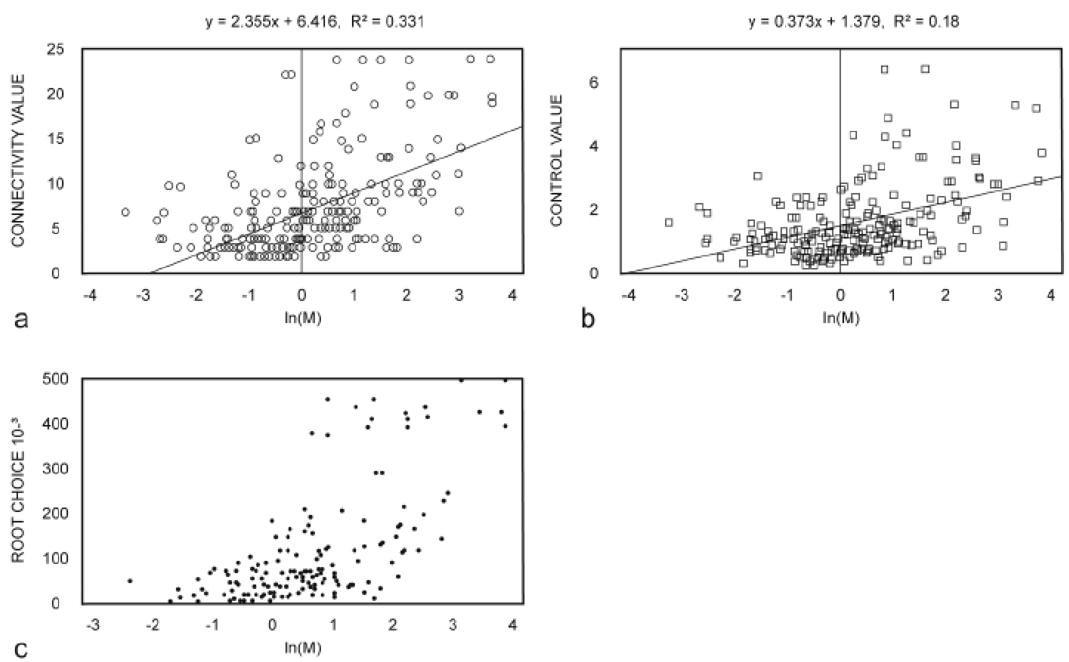


Figure 14.12 Plots of (a) connectivity values, (b) control values and (c) global choice values against the natural logarithm of movement rates (M).

the pattern of integration, derived from a purely configurational analysis of the grid, and the pattern of real movement is both strong and pervasive.

But why, we might ask, is the relation logarithmic? We can throw light on this and explore the data further by breaking the data down into its 10 subareas. Table 14.2 tabulates key results from the area-by-area analysis, most notably the

Table 14.2 Mean spatial values, encounter rates and correlation coefficients for 10 subareas of King's Cross

	Subarea[a]									
	1	2	3	4	5	6	7	8	9	10
Integration of system										
1,211 observations [b]	0.63	0.66	0.65	0.69	0.67	0.65	0.67	0.92	1.14	0.88
796 observations[c]	0.65	0.72		0.71	0.62	0.66	0.69			
Mean connectivity										
1,211 observations[b]	7.63	8.07	5.59	8.31	9.14	8.06	8.74	5.22	4.86	3.75
796 observations[c]	6.86	5.46		7.68	9.57	6.12	7.63			
Maximum connectivity										
1,211 observations[b]	19	21	21	24	24	20	20	10	7	6
Encounter rates										
Moving adults (phm/min)										
with shopping streets	4.22	1.58	1.81	1.58	5.47	4.80	9.02	0.27	0.31	0.57
without shopping streets	2.71	1.58	1.81	1.58	3.25	2.68	3.39	0.27	0.31	0.57
men	2.76	0.93	1.02	0.78	2.88	2.97	5.77	0.13	0.14	0.27
women	1.46	0.65	0.78	0.80	2.58	1.83	3.25	0.14	0.17	0.30
Stationary adults (phm)										
men	0.71	0.78	0.64	0.17	0.92	1.10	1.35	0.06	0.12	0.43
women	0.30	0.33	0.25	0.10	0.95	0.63	0.81	0.04	0.08	0.07
all adults	5.23	2.69	2.70	1.85	7.34	6.53	11.18	0.37	0.51	1.06
children	0.20	0.30	0.22	0.23	0.28	0.19	0.21	0.88	0.73	0.44
Pearson's *r* correlation coefficient between integration and										
log of moving adults	0.71	0.70	0.25	0.67	0.78	0.64	0.80	0.68	0.70	0.62
moving adults[d]	0.54	0.56	0.25	0.70	0.55	0.51	0.79	0.74	0.74	0.67
without shopping streets[d]	0.73	0.72	0.41	0.72	0.70	0.66	0.82			

[a] 1 Southeast area, 2 East area, 3 Northeast area, 4 North area, 5 West area, 6 Southwest area, 7 South area, 8 Maiden Lane estate, 9 Elm Village estate, 10 Bemerton estate.
[b] With housing estates.
[c] Without housing estates.
[d] Not log transformed.

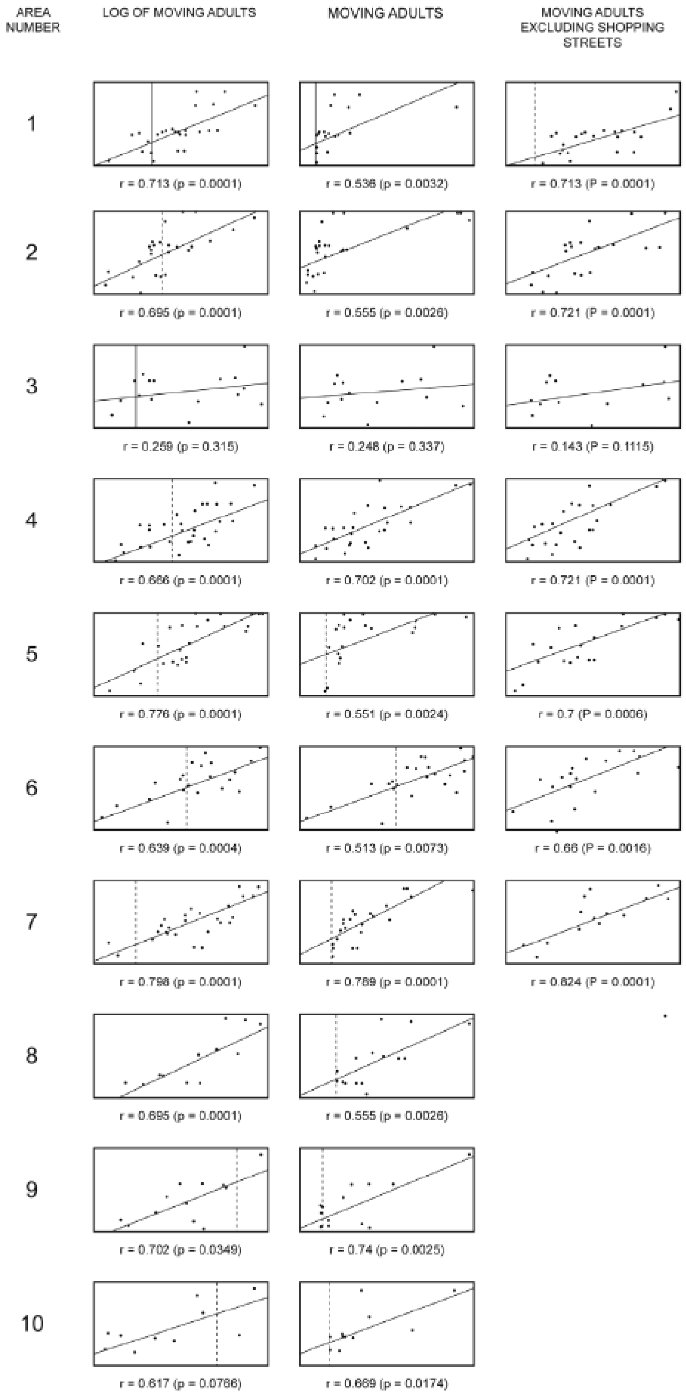


Figure 14.13 Scattergrams plotting integration values against logged moving adults, moving adults and moving adults (excluding shopping streets) for 10 subareas of King's Cross.

correlation between integration and the logarithm of movement rates for each area, the correlation between integration and unlogged movement rates, and the correlation between integration and unlogged movement for each of the areas excluding their shopping streets (that is, their main attractor spaces). Figure 14.13 shows scattergrams for each of these three relations for all areas. In all cases, integration values are still read from the largest system.

These data show certain remarkable consistencies. The first column shows that in all cases, except the Northeast area (3), correlations between integration and logged movement rates are strong and highly significant, in spite of the fact that integration is still being read from the largest area. In the case of the Northeast area, it is clear that the area does not operate as a single area, but as two subareas, one residential, the other light industrial, linked and separated by a major road.

The scattergram reflects this split.

The second column shows that when movement rates are unlogged, a number of spaces appear well above the regression line. These spaces, it turns out, are in all cases the spaces with groups of shops. The third column shows that, when plotted without these spaces, the relation between integration and movement becomes directly linear. In the case of the North area (4), in which the route observed has only one line with a few shops, we find that the relation is already, more or less, linear. Even then, the elimination of the shopping space (which in this case does not have the highest encounter rate) improves the correlation quite significantly. Last, in all three housing estates, where there are no groups of shops (in area 10 there is one shop which has no measurable effect on the movement pattern) the unlogged encounter rate produces substantially stronger correlations than the logged version.

The inference is unavoidable. The effect of shops as attractors is to shift a basically linear relation between integration and movement into a logarithmic relation. It might not be going too far to suggest that the consistency of the logarithmic effect suggests that each of the shopping spaces has about the right number of shops for its integration value. Shops, as the basic attractors in urban areas, work, it seems, as logarithmic multipliers on a basic pattern of movement

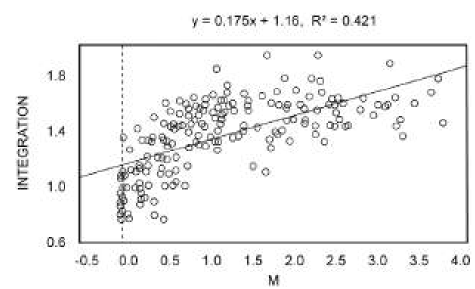


Figure 14.14 Plot of integration against numbers of moving adults (M) in all non-shopping streets.

defined by configuration. By and large, they appear to have been located in the right spaces and in numbers proportionate to their integration value.

But the matter does not quite end there. If we take the whole data set and eliminate the same shopping spaces, as we did in the areas singly, the relationship between integration and movement rates remains markedly logarithmic (see Figure 14.14). As the elimination of shopping spaces does eliminate the logarithmic effect in the areas taken on their own, the only possible inference is that there is also a global, logarithmic effect at the level of relations among the areas. This can be shown by plotting integration against unlogged movement for each area (except area 3) on the scale defined by the scattergram for the whole system (as shown in Figure 14.11a). By setting the areas, more or less, in order of their movement rates, it becomes clear that as mean rates increase, so there is a strong tendency for the regression line for the area to shift from near vertical to near horizontal. In other words, the distribution of grid properties and movement rates at the area level is, itself, logarithmic (see Figure 14.15).

Unfortunately, with the present database and software, it is not possible to ascertain how far this is a result of grid properties and how it is a result of attractors. For example, it is not yet possible to excise areas from the larger system and compute their mean integration values as computed, or read, within the larger system in order to then compare them with mean encounter rates. Nor is it possible to take the 10 routes as representative of their well-defined areas. For example, in the Northeast area, the grid is sparse and block sizes are large, with the effect that the observation route is dominated by the main grid, with few smaller-scale spaces.

The range of integration values for observed lines is consequently less, and the average integration value does not reflect the weakness of the area in the distribution of integration in the large area. In contrast, the West area includes, not only the major shopping attractor of Camden High Street, but also the strongly segregated region to the immediate west of the King's Cross site. It therefore includes both strong integration and strong segregation, giving it a high range of integration values, and also shows the influence of both positive and negative attractors.

However, it does seem that, over and above the effect of specific attractors in the area, there is also a global pattern of what we might call attractor areas, and that these reflect both global properties of the grid and differences in the distribution of attractors. It seems likely, but we cannot be sure, that the latter is consequent on the former. What can be said is that the global pattern of area differences in movement rates follows the distribution of the integration core of the whole area, as discussed earlier in this paper, singularly closely.

What about the 'negative attractors'? Are they also basically configurational? In the case of the falloff effect of the King's Cross site, the site is both a configurational lacuna and also has a radical absence of attractors. Intuitively, in view of the powerful relation between spatial configuration and movement, it seems unlikely that a major interruption of route continuity and destination availability in a substantial area would not have local 'negative multiplier' effects on the surrounding area.

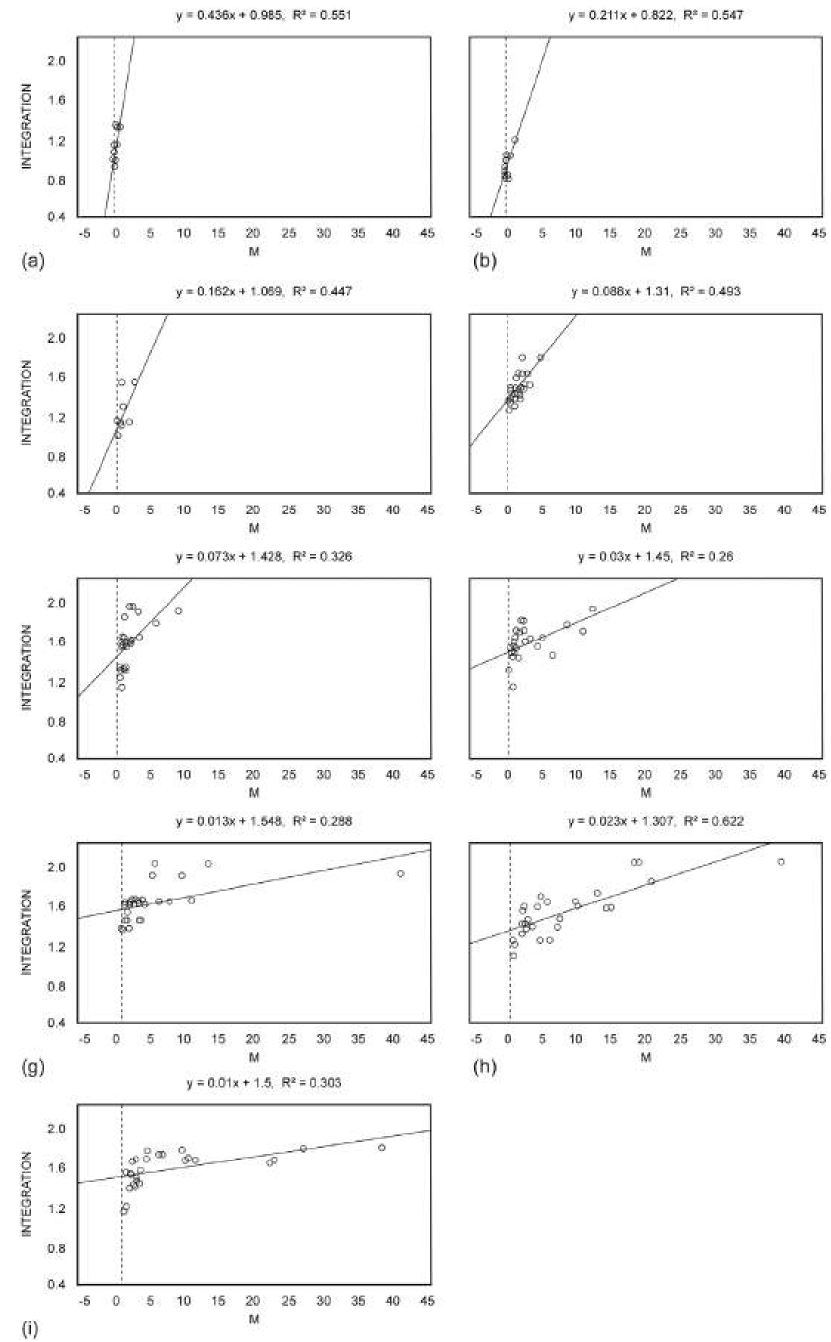


Figure 14.15 Plot of integration against numbers of moving adults (M) in (a) Maiden Lane estate; (b) Elm Village estate; (c) Bemerton estate; (d) North area; (e) East area; (f) Southwest area; (g) Southeast area; (h) South area; (i) West area.

This, in turn, could be expected to have a knock-on, negative effect on attractors in the vicinity of the site. However, it is difficult to show this unambiguously and numerically. This would seem to be an area where much further research is needed. On present evidence, it seems only a reasonable inference that the almost uniform blight that seems to characterise the area surrounding the site is a long-term, configurational effect.

In the case of the housing estates, however, it is possible to show that the extraordinarily low movement rates are a configurational effect. First, we can eliminate one obvious possibility: that of housing density. On the whole, the estates have higher densities, in terms of population, floor space, and usually ground coverage, than their urban surroundings. Nor is it easy to see how the lack of groups of shops could lead to such an extreme diminution of movement rates. There is, it seems, no simple attractor explanation.

In fact, it is easy to use the axial map to show that the reduction of movement rates results from configuration, not attraction. Figure 14.16 is a '10-minute map' of the movement pattern on the observed route in one of the three housing estates, the Maiden Lane estate immediately to the north of the King's Cross site. Each dot represents one moving adult encountered during an average 10-minute period throughout the day (that is, it multiplies the phm/m movement rate by 10 to give graphical clarity).

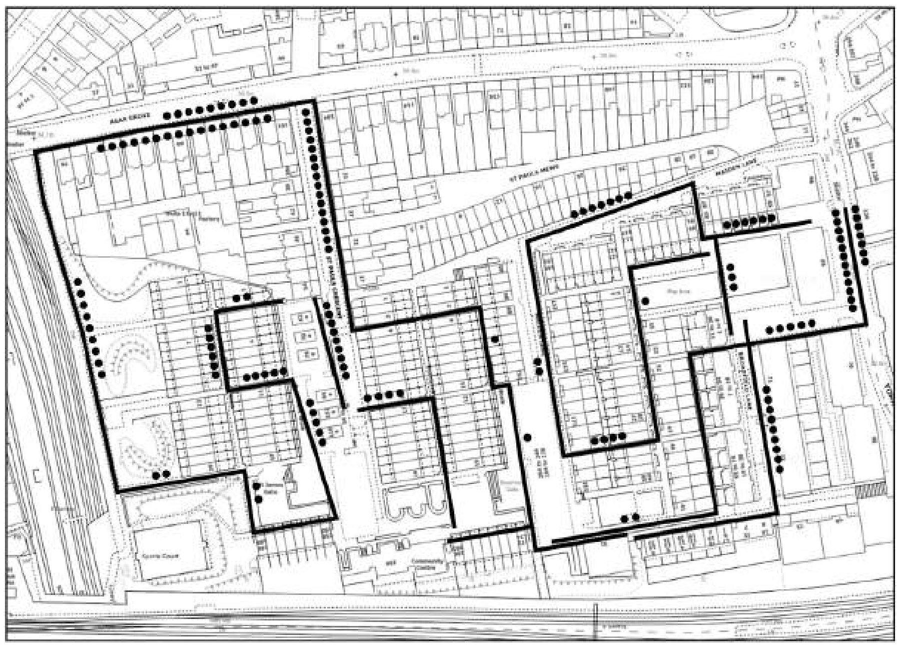


Figure 14.16 '10-minute map' of the movement pattern by adults on the observed route in the Maiden Lane estate.

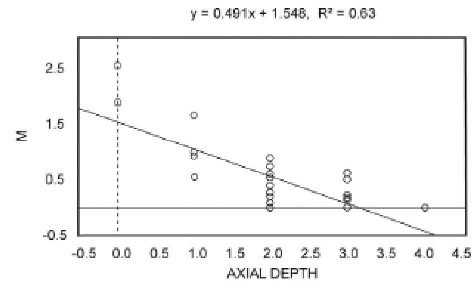


Figure 14.17 The negative relation between axial depth from the exterior and movement rates (M) ($r = 0.793$, $p < 0.000$).

The map shows that, on the lines peripheral to the estate, movement rates approximate the norm for the area. On the lines entering the estate from the periphery – one axial step deep – rates are considerably lower, although still significant. At two axial steps into the estate, rates drop again, then again at three steps.

This negative relation between axial depth and movement rates can be shown as a scattergram (Figure 14.17) and indexed by a correlation coefficient (in this case $r = 0.793$, probability $p < 0.0001$). This rapid falloff in movement with depth is a very common property for housing estates and shows, graphically, how spatial configuration freezes out all natural movement from the estate.

The configurational influence of spatial segregation on movement within the estates, can be confirmed by looking at average integration values and movement rates for the estates and contrasting them to the street areas. We find that the mean integration value of observed lines in the urban areas is 1.508, very close to the average for London streets as a whole, with a minimum of 1.045 and a maximum of 2.004. The mean movement rate is 3.943 phm/min, or 2.982 phm/min without main shopping streets. In the estates, the mean integration is 1.020, that is, well below the *minimum* for the street areas, and the maximum is 1.522, or about the mean for street areas. The mean movement rate is 0.365 phm/min.

It is not possible to tabulate mean integration against average movement rates for all of the areas, because current technology does not permit the excision of a group of connected spaces while still reading their integration values from the larger system. To use only the observed spaces in each area would entail the risk that the sample of spaces does not truly represent the area. We can, however, divide the whole data set into a number of integration bands with equal numbers of spaces and check the average integration of these against mean movement rates. Figure 14.18 shows this for six integration bands, yielding a correlation of $r = 0.99$.

The effect of low movement rates on the awareness of others can be further clarified by setting rates against the mean length of axial lines in areas to derive a figure for how much time a pedestrian spends in and out of visual contact with other people. For urban residential streets, a mean natural movement rate of

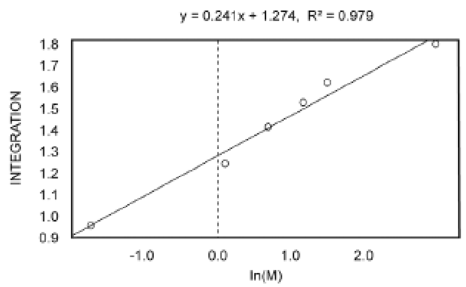


Figure 14.18 The average integration of six equal-sized groups of spaces against mean movement rates (M) for the whole dataset ($r = 0.99$, $p < 0.0001$).

2.7 phm/min, coupled to a mean length of axial line of about 300 m, ensures that a pedestrian is in virtually constant visual contact with others, and usually with several. A mean rate of 0.27 phm/min, and a mean axial length of 50 m (as found on this estate) would mean that a pedestrian would be *out* of visual contact for about 88 per cent of the time, a remarkable reversal of urban norms. As we have observed elsewhere, many housing estates have lower degrees of awareness of others through movement in the middle of the day than even quiet residential areas have at midnight. Coupled to the scaling down of space, which is an almost invariable property of such estates, one begins to arrive at a numerical picture of the extreme sense of isolation which people often report in modern housing estates.

Case study 2: The City of London

These inferences may be explored further by looking at a study of the City of London, carried out for the Mansion House Square public inquiry in 1984 to try to establish whether or not the proposed square would work – and leading, incidentally, to the prediction that it would work very well. Again, this review deals only with the analytic aspects of the study relevant to natural movement.

Figure 14.19 is a plan of the City of London 'within the walls', Figure 14.20 is a black-on-white space map, and Figure 14.21 is its axial map. Figure 14.22 shows the three routes observed during the study, and Table 14.3 is a tabulation of the movement rates for each of the five time periods for the two central routes with averages for all time periods taken together, for the morning and afternoon taken together, and for the two rush-hour periods taken together. The third route is omitted both because it was primarily concerned with the modern development around St Paul's rather than with the urban fabric of the city, and because it is too close to the western periphery to the City to be comparable with the other two

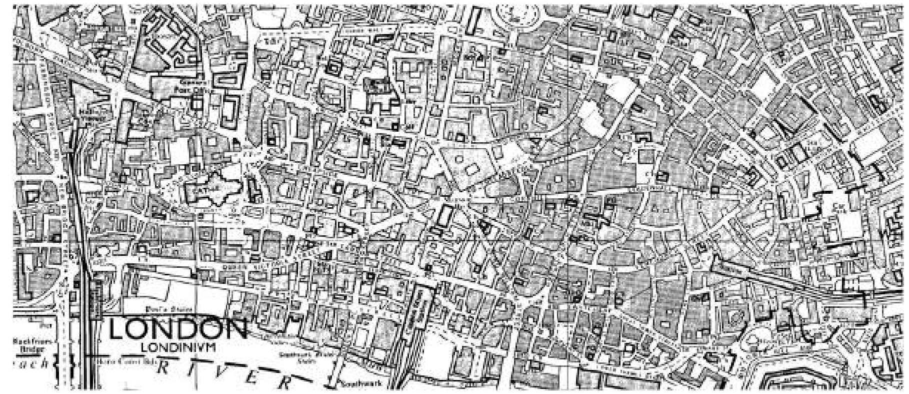


Figure 14.19 Map of City of London 'within the walls'. National Grid 1:10,000, circa 1993. Historic, using EDINA Historic Digimap Service.

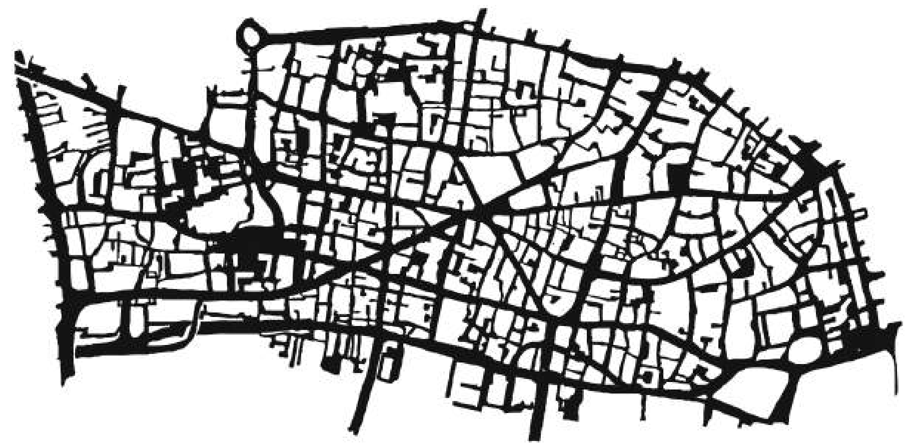


Figure 14.20 Black-on-white space map of the City of London.

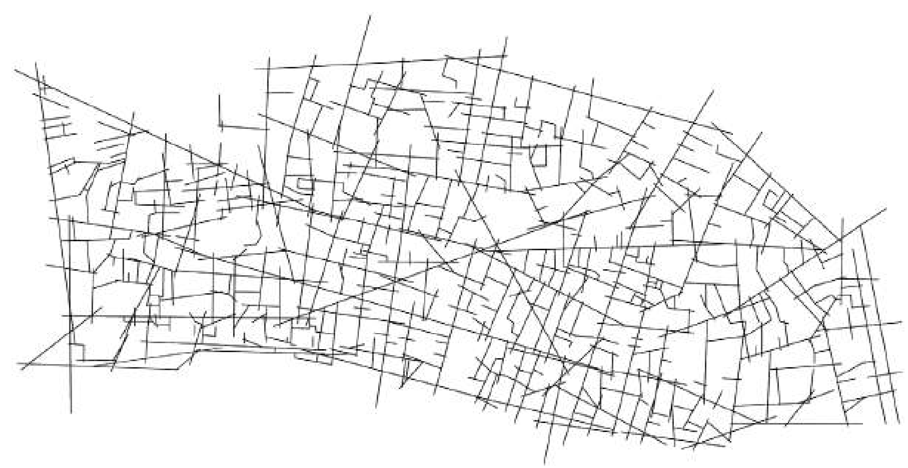


Figure 14.21 Axial map of the City of London 'within the walls'.

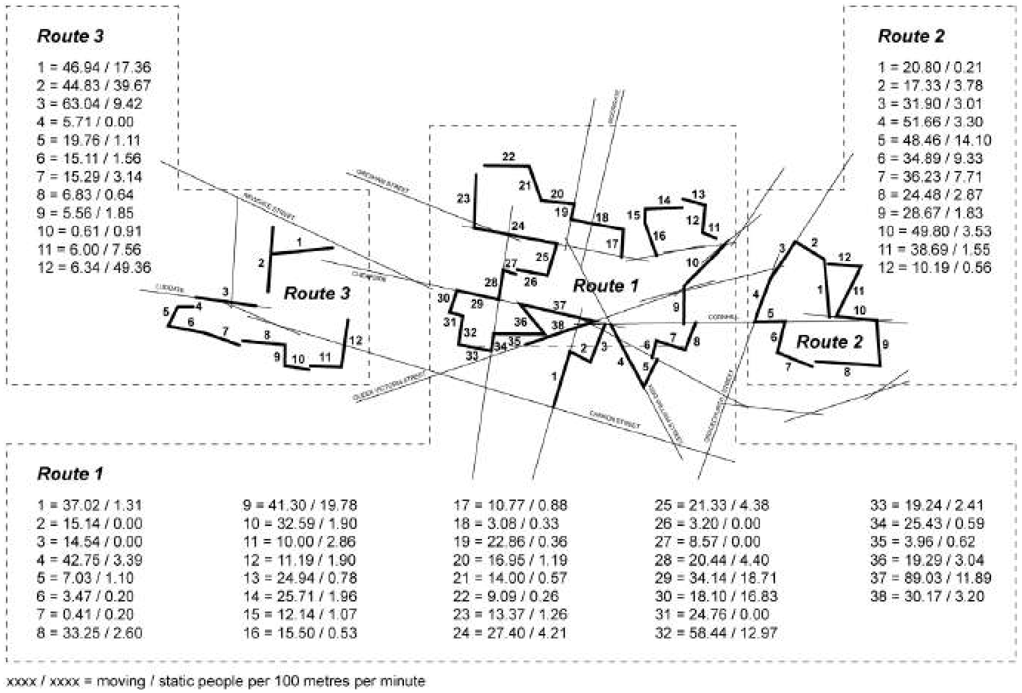


Figure 14.22 Three observed routes annotated through the City of London, with density of moving and static people during the midday period (12–2pm).

routes. Because it was shown that the distribution of rates was very similar for the morning, afternoon and midday periods (though the last had, of course, much higher rates), and contrasted to the rush-hour periods when the distribution was different in quite marked ways, the figures used below will be the average for the three middle periods of the day.

Figure 14.23a is the scattergram for integration values plotted against movement rates in the two routes. Although the correlation is respectable, the scattergram is decidedly untidy, and appears to contain sets of points suggesting different regression lines. It turns out, on closer examination, that these sets of points are connected lines forming subareas. We can explore matters further by breaking up the scattergram into its component subareas, but holding the axes of Figure 14.23a steady so that the different areas can be seen, within the same frame. Figure 14.23b is the 'central' area on its own, that is, the area to the south of Gresham Street, with all lines to the north of Gresham Street eliminated along with the whole of the eastern route. The correlation for this more homogeneous set of connected lines becomes very much stronger.

We then take the two subareas north of Gresham Street and plot them separately in Figures 14.23c and 14.23d. We see that each has its own characteristic

Table 14.3 City of London movement rates for each of the five time periods for the two central routes (route 1 and route 2)

	Time period average						AM rush hour	PM rush hour
	1	2	3	4	5	all		
Route 1 (moving people per 100 m)								
1. Walbrook	86.1	17.8	37.0	18.3	110.3	53.9	18.1	98.2
2. St Stefen's Row	9.3	4.3	15.1	5.4	10.0	8.8	4.9	9.7
3. Mansion House Place	15.8	4.5	14.5	6.8	17.3	11.8	5.6	16.5
4. King William Street	32.6	24.4	42.8	23.0	33.8	31.3	23.7	33.2
5. Abchurch Lane	9.0	4.1	7.0	5.5	8.5	6.8	4.8	8.7
6. Cornhill, Lombard Street Alley	2.1	1.2	3.5	2.4	5.5	2.9	1.8	3.8
7. Change Alley	1.0	1.4	0.4	0.6	1.2	0.9	1.0	1.1
8. Birchin Lane	18.8	14.5	33.2	14.3	14.2	19.0	14.4	16.5
9. Royal Exchange Building	40.3	20.0	41.3	23.3	41.0	33.2	21.6	40.6
10. Old Broad Street	38.1	18.3	32.6	14.6	37.8	28.3	16.5	37.9
11. Austin Friars	9.4	5.6	10.0	8.1	8.3	8.3	6.8	8.9
12. Austin Friars	6.7	6.7	11.2	6.7	5.8	7.4	6.7	6.2
13. Austin Friars	12.4	20.6	24.9	9.6	7.0	14.9	15.1	9.7
14. Copthall Building	12.1	20.2	25.7	14.8	11.9	16.9	17.5	12.0
15. Angel Court	7.9	11.2	12.1	6.1	5.0	8.5	8.7	6.5
16. Angel Court	13.5	11.2	15.5	11.8	13.8	13.2	11.5	13.0
17. Tokenhouse yard	7.2	4.6	10.8	7.7	8.7	7.8	6.2	7.9
18. King's Arms yard	3.1	3.1	3.1	3.3	4.5	3.4	3.2	3.8
19. Coleman Street	30.0	12.9	22.9	5.0	40.4	22.2	9.0	35.2
20. Masons Avenue	9.1	8.0	17.0	9.0	7.6	10.1	8.5	8.3
21. Basinghall Street	9.8	5.2	14.0	7.9	15.0	10.4	6.5	12.4
22. Space behind Guildhall	10.5	4.5	9.1	4.3	14.7	8.6	4.4	12.6
23. Aldenmanbury	10.1	8.1	13.4	8.5	19.3	11.9	8.3	14.7
24. Gresham Street	12.7	14.1	27.4	14.4	29.8	19.7	14.3	21.2
25. Old Jewry	17.8	10.0	21.3	14.9	34.4	19.7	12.4	26.1
26. St Olave's Court	0.4	1.6	3.2	0.8	2.2	1.6	1.2	1.3

distribution and its own regression line. Figure 14.23d, the west complex (around the Guildhall) has a narrow range of movement values, but a wide range of integration values, giving a very strong correlation coefficient, logged or unlogged (there are no shops in this area). The eastern complex of the northern area (Figure 14.23c, Austin Friars) shows narrow ranges both in natural movement

	Time period average						AM rush hour	PM rush hour
	1	2	3	4	5	all		
27. Prudent Passage	3.3	7.8	8.6	0.0	1.1	4.2	3.9	2.2
28. King Street	14.4	11.5	20.4	12.5	26.9	17.2	12.0	20.6
29. Cheapside	54.8	45.0	134.1	38.9	65.3	67.6	41.9	60.1
30. Bow Churchyard	14.1	8.1	18.1	10.2	24.1	14.9	9.2	19.1
31. Bow Churchyard Alley	15.6	11.1	24.8	7.6	24.4	16.7	9.4	20.0
32. Bow Lane	24.2	17.9	58.4	22.1	42.7	33.1	20.0	33.5
33. Watling Street	11.8	12.4	19.2	11.0	29.1	16.7	11.7	20.4
34. Queen Street	14.0	14.7	25.4	9.4	12.3	15.2	12.0	13.2
35. Pancras Lane	1.6	2.5	4.0	2.9	4.2	3.0	2.7	2.9
36. Bucklersbury	45.6	7.7	19.3	16.1	65.6	30.9	11.9	55.6
37. Cheapside	43.1	36.1	89.0	45.6	71.9	57.1	40.9	57.5
38. Queen Victoria Street	27.9	19.3	30.2	20.7	30.8	25.8	20.0	29.3
Route 2 (moving people per 100 m)								
1. P&O and Commercial Union Square	31.2	7.9	20.8	8.5	32.5	20.2	8.2	31.9
2. Great St Helen's	30.3	11.1	17.3	10.2	27.1	19.2	10.7	28.7
3. Bishopsgate	57.6	15.6	31.9	14.3	73.5	38.6	14.9	65.6
4. Gracechurch Street	67.0	26.2	51.7	24.3	68.3	47.5	25.2	67.7
5. Cornhill	43.5	36.4	48.5	19.2	44.1	38.3	27.8	43.8
6. Whittington Avenue	22.0	18.9	34.9	22.7	16.9	23.1	20.8	19.4
7. Leadenhall Place	21.2	19.4	36.2	19.6	22.9	23.9	19.5	22.1
8. Fenchurch Avenue	28.1	19.1	24.5	21.0	19.5	22.4	20.1	23.8
9. Billiter Street	34.0	17.2	28.7	16.7	58.5	31.0	16.9	46.2
10. Leadenhall Street	68.2	24.5	49.8	19.2	66.3	45.6	21.9	67.3
11. St Mary Axe	38.0	18.4	38.7	13.2	33.7	28.4	15.8	35.9
12. Undershaft	6.1	4.3	10.2	4.4	4.6	5.9	4.4	5.4

[a] Average over all time periods, average over morning period 2 and afternoon period 4 (am/pm), and average over two rush-hour periods.

rates and in integration, and clusters in a very small region of the scattergram, though retaining a reasonable correlation.

To turn to the eastern route, Figure 14.23e is a plot of the scattergram (still using the same coordinates for comparability) and, again, shows a distinct regression line and a good correlation. But this area falls on two sides of the

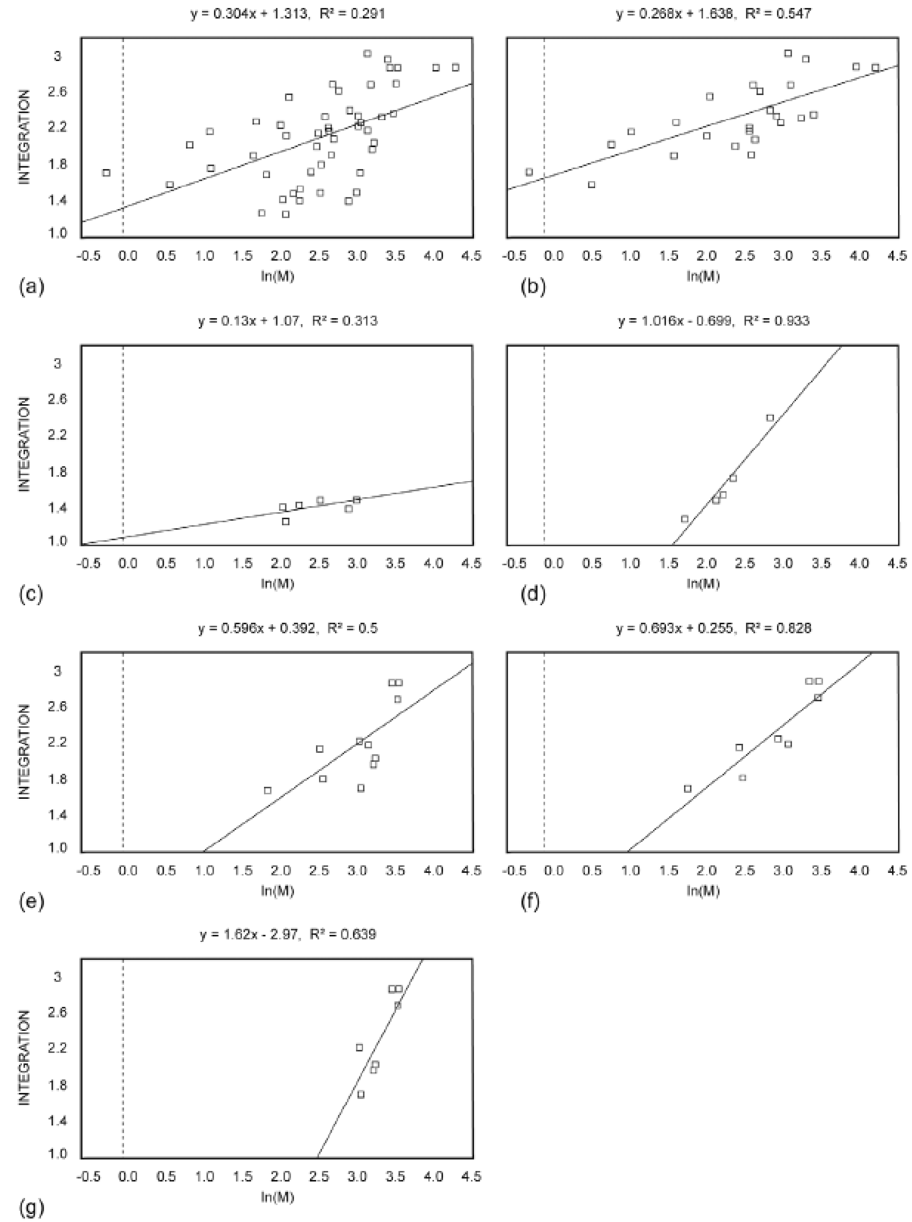


Figure 14.23 Scattergrams plotting integration values against logged movement rates (M) for (a) all lines in the two routes; (b) the 'central' area on its own; (c) the eastern subarea to the north of Gresham Street around Austin Friars; (d) the western subarea to the north of Gresham Street and around the Guildhall; (e) the eastern route; (f) the eastern route subarea with the much quieter area of St Mary Axe to the north; (g) the eastern route subarea with Leadenhall Market to the south.

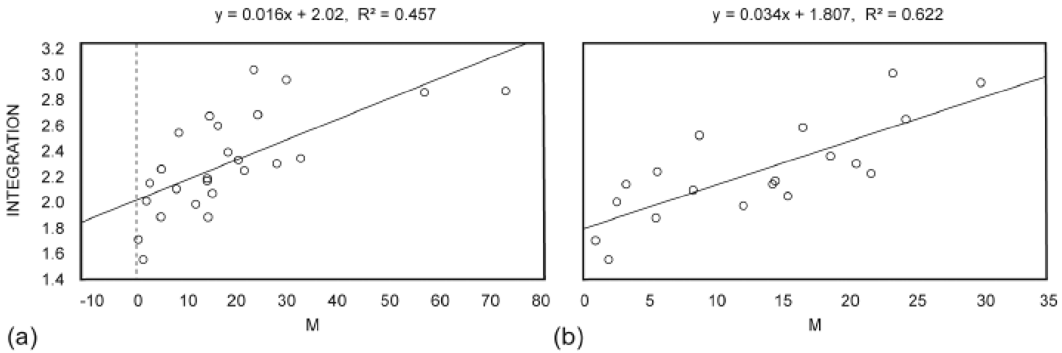


Figure 14.24 Plot of integration against natural movement rates (M) unlogged (a) with and (b) without shopping streets.

Cornhill-Leadenhall Street line, with Leadenhall Market to the south and the much quieter area of St Mary Axe to the north. Figures 14.23f and 14.23g show the scattergrams for these areas separately, although each time the Cornhill-Leadenhall Street line is included. Once again, one can see that the effect of collections of shops is to distort the regression line upwards, but still to conserve the correlation with integration.

The logarithm hypothesis about lines with shops can be further explored by looking again at the central area. Figures 14.24a and 14.24b show integration against natural movement rates unlogged, first with, then without, shopping streets. In Figure 14.24a, with all the data still present, the log effect is easily seen, whereas in Figure 24b the scattergram has become linearised. The logarithmic effect of shops seems again clear.

The City case study thus shows a rather different kind of relation between integration and movement, one in which differences between local subareas seem strong. These differences may be the result of a differential distribution of built forms (for example, the higher than expected rates in the Austin Friars area could be caused by the presence of tall buildings), or to the distinctively regional character of the City grid.[12] Further research would be needed to clarify this. But of the strength of the basic relation between the pattern of integration and movement, there need be no doubt.

Case study 3: A South London housing estate

Further aspects of the relation between integration and natural movement may be explored by looking at the relevant parts of a study of the relations between spatial patterns and crime, recently carried out in collaboration with the Crime Prevention Unit of the Home Office. The study of natural movement, in and around

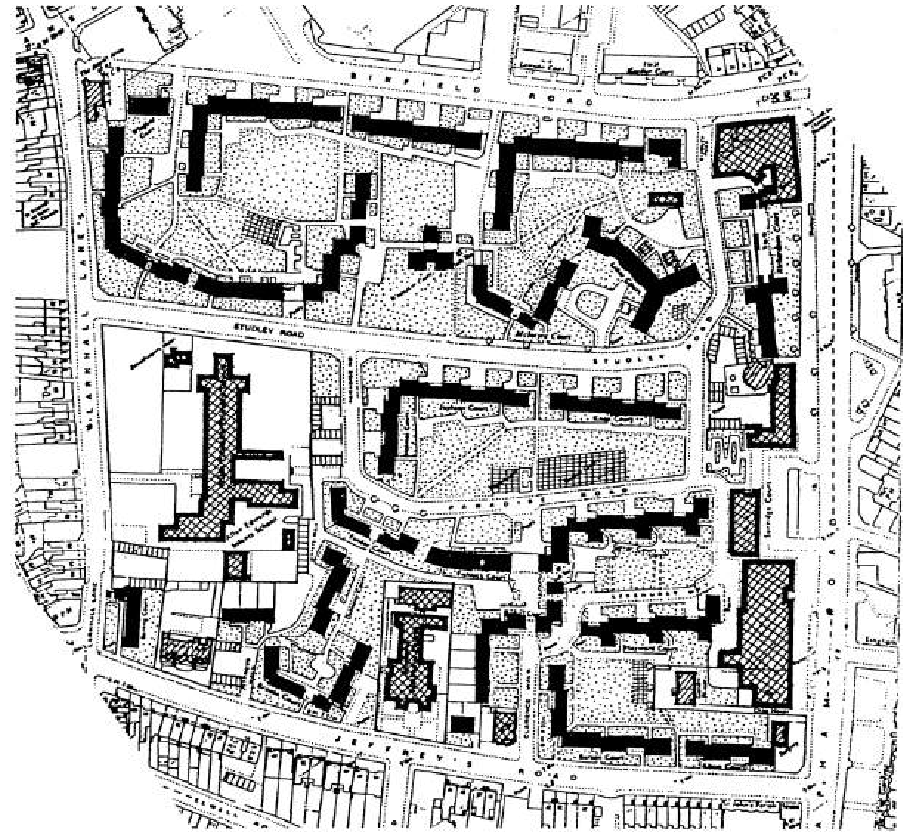


Figure 14.25 Plan of a South London housing estate.

the estate, was part of the study because it seemed possible that it might be an intervening variable between space and the vulnerability of locations to different kinds of crime.

Figure 14.25 is the map of the estate, Figure 14.26 its black-on-white map in its surrounding urban area, and Figure 14.27 its axial map. Four different levels of representation were used for calculating spatial values: the estate only without its surrounding streets (measures suffixed 1); the estate with its surrounding streets (suffixed 2); the estate plus the street system in the surrounding urban area as shown in Figure 14.26, but without the housing estates (suffixed 3); and the last, plus other housing estates in the area (suffixed 4).

Of particular interest here is the effect of reading integration values from the four different systems in post-dicting the pattern of movement. Figures 14.28a–d show scattergrams for movement rates using the four systems. The best representation is the third, the estate plus its urban surroundings, which gives a correlation coefficient of $r = 0.752$.

Figure 14.26 'Black-on-white' map of the estate in its urban context (including other housing estates).

Dividing the data into the surrounding area, Figures 14.29a and 14.29b, and the estate, Figure 14.29c (using this reference system), and using the untransformed movement rates, we find that the surrounding area is strongly logarithmic, but the estate itself is not. The correlation coefficient for the surrounding area is also a good deal better than for the estate interior, even though this estate is relatively open and permeable to its surroundings. Once again, we find that movement rates fall rapidly with depth from the outside and, indeed, we again find depth from the outside strongly negatively correlated with integration.

Figure 14.27 Axial map of the estate and its urban context (including all other housing estates).

The relation of these patterns to the spatial distribution of crime on the estate has been explored in the remainder of this study. A series of strong – though occasionally complex – relationships have been found between spatial patterns and crime. But these are not our theme here. The results are now being written up separately, and are currently available as a working paper.

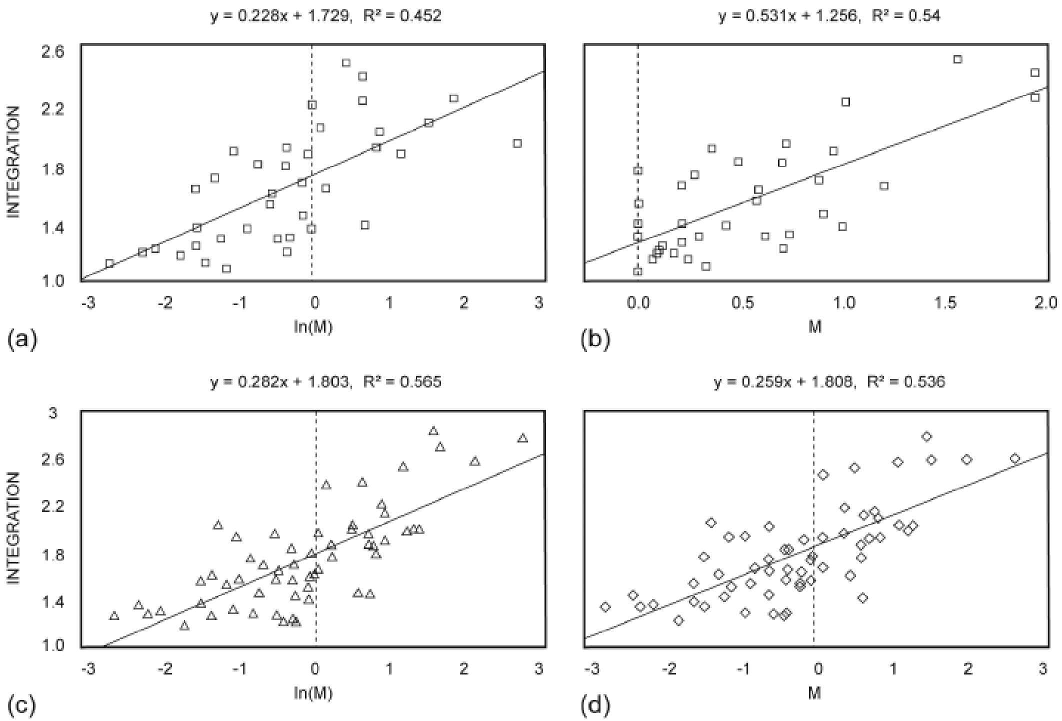


Figure 14.28 Scattergrams for integration values against movement rates (M) calculated in (a) system 1; (b) system 2; (c) system 3; (d) system 4.

Other studies

A considerable body of other studies now exists, some published, others as aspects of applied research or PhD theses. A number are worth referring to here. First, a study of Highgate and its contextual area was carried out for the Islington Health Authority as part of a development initiative for a major site in the area. The study was of particular interest because Highgate is essentially a single main street on an important through route, but surrounded by an area with a rather suburban character in that the islands are very large, and the buildings are relatively low density.

Two scattergrams are sufficient for our purposes here. Figure 14.30a shows the relation between integration and unlogged movement rates, and Figure 14.30b the same relation, taking the logarithm of movement ($r = 0.818$, $p = 0.000$). The highest movement spaces are the three observed sections of the High Street. The logarithmic effects of the attractors here closely follow those of the previous cases, and without them the relation is linearised.

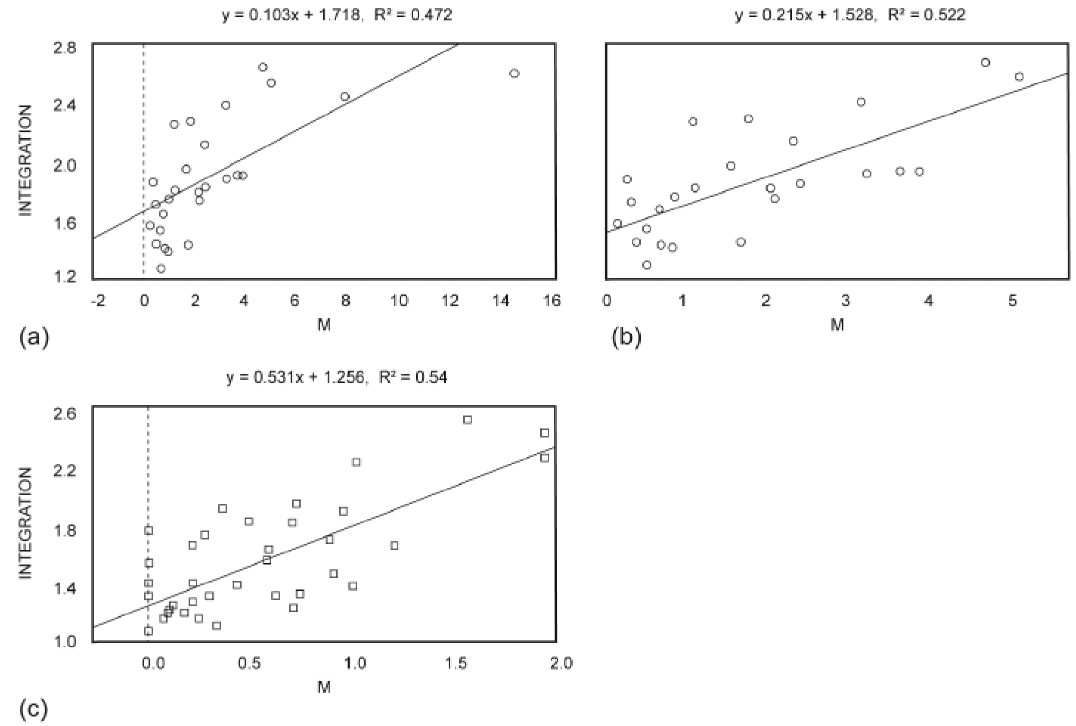


Figure 14.29 (a) Scattergram of integration in system 3 against untransformed movement rates (M) of the surrounding area; (b) with the two attractor spaces excluded; and (c) for the estate itself.

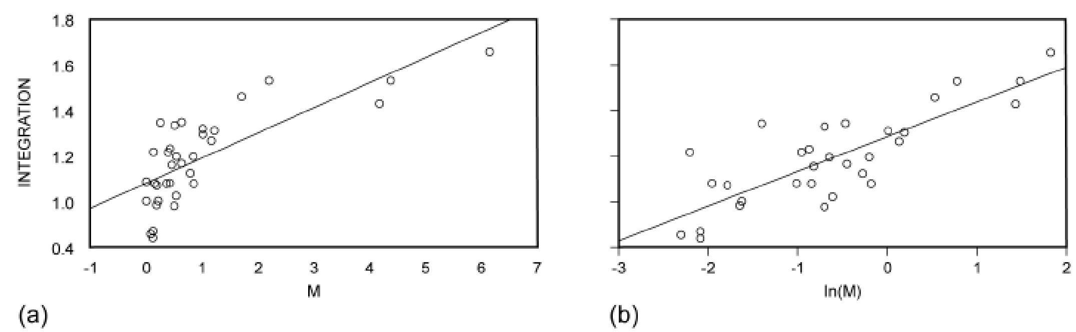


Figure 14.30 The relation between integration and (a) unlogged and (b) logarithm of movement rates in the Highgate study area.

Previously published cases include 'Creating life: Or, does architecture determine anything?',[13] in which a series of studies of urban areas, suburban areas and housing estates are reported, in all cases with strong and significant relations between integration and movement. In this paper, a number of cases are reported where integration works less directly, or less well, in post-dicting movement, though still, in all cases, reaching high levels of statistical significance. It is hypothesised that this lower performance is caused by the absence of a 'second-order' configurational property called 'intelligibility', defined as the degree of correlation between the connectivity of lines and their integration value. That is, between what can be seen of the line visually and locally, and how this relates to the importance of the line in the system as a whole. This hypothesis is currently being explored in theoretical studies being carried out in the Unit for Architectural Studies.

Also recently published is a set of studies of six Greek towns by Peponis et al. in an issue of *Ekistics* devoted to space syntax research.[14] All show strong correlations between spatial parameters and observed movement, with integration, again, the strongest performer. This study is unusual in the degree to which the detailed morphology of the towns is considered and the effects of, for example, the opening and closing of shops. Xu Jianming, in a doctoral study, has found strong correlations in all nine housing estates studied, although, again, with variations which, he hypothesises, could be the result of the degree to which 'intelligibility' as defined above is present in the system. Krüger reports (personal communication) good correlations in all but one of a set of studies of urban areas in Brazilian towns. In the nonconforming case, the study was based on a single round of observations (the norm is 20–30) by undergraduate students. Inspection of maps suggests, however, that the contextual area for the study was insufficient to give an adequate syntactic picture of the study area.

Apart from Krüger's anomaly, no cases are known where sufficient analysis and observation has not led to a significant relation of some kind between integration and movement. The proposition that the results of such studies as these permits us to offer a theory of *natural movement*. Namely, movement which is, ceteris paribus – other things being equal – determined by the configuration of the grid, is sufficiently supported by a range of studies to suggest that configuration ought to be given, at least, as equal weight as attraction in the design of urban areas. Cases are not standard, however, and much remains to be understood about the *ceteris paribus* phrase.

Implication for theory: Why axiality? Why linearity?

If such structures as integrating cores are as fundamental to urban form as the results of natural movement studies suggest, then there are implications for theory which go beyond the analytic theories we need in research, and into the normative theories we use in design. Recent debates about urban space have been much preoccupied with two theoretical questions: which urban *scale* is the 'human scale' and how do urban forms become *intelligible*? Most commonly, answers are sought in some interpretation of historical precedent. The scale question leads to a proposal to reproduce the allegedly smaller-scale environments of the past, and the intelligibility question to a proposal to introduce historically meaningful monuments and landmarks to act as orientational devices in the urban landscape.

The implication of natural movement studies would seem to be that both scale and intelligibility are in the first instance questions of the morphology of space and cannot be separated from the spatial configuration of the grid itself. There is also a clear implication that there has been a substantial misreading of urban morphological history. Human scale in urban space seems to be more about ensuring that scale is sufficiently large rather than small, and intelligibility seems to arise in the first instance as a product of the same upwards scale effects, rather than from physical cues and symbols – though these undoubtedly also play some role.

These issues can, we believe, be clarified by posing two questions implied by our results: first, why axiality? That is, why should the axial organisation of space be so fundamental? And, second, why linearity? That is, why should the relation between spatial and functional patterns be a simple linear one? To answer these questions, we must depart for a moment from the formal and statistical analysis of spatial and functional patterns and consider urban space and form from the point of view of the individual subject who experiences it and uses it. If axiality and linearity are key morphological properties of urban space and function, then there must be some sense in which both 'get into the head' of individuals and come out as behaviour.

On the first question, 'why axiality?', kinds of data converge: first, the effects of descaling modern space, especially residential space, which we looked at in the King's Cross study, and, second, data on syntactic aspects of simulated urban form generation. On the first, we may turn back to Figure 14.7 and remind ourselves that a prime spatial attribute of modern housing estates is the *reduction* of spatial scale and the increase in spatial complexity. This is as true of the medium-rise estates to the west of the Caledonian Road as it is of low-rise estates to the north and west of the King's Cross site. It is also as true of geometrically ordered designs as of more organic-seeming layouts. Much analytic experience has convinced us that both of these are generic properties of modern layouts.

As shown in 'Creating life',[15] and since confirmed by further studies, the functional effects of this down-scaling of space are usually that the integration core is peripheralised, so that the estate interior lacks structure; that, consequently, movement is peripheralised, and people simply move out of the estate by the shortest possible route; that a dramatic falloff in movement rates occurs with depth, so much so that it becomes normal for people to find themselves on their own for most of the time when moving about within an estate; and that probabilistic social interfaces are disrupted. The net result is that distinctive sense of isolation amidst high densities of people that characterises so many housing estates. Extensive studies of such environments have persuaded us that the problem of spatial scale in modern design is not the overscaling of space, but its radical underscaling.

We may come at this issue from another angle by looking at scale in urban form generally. It has been noted in previous texts commenting on attempts to simulate urban spatial properties through partially randomised cell-aggregation processes that local rules for growth that work for small aggregations do not work for larger aggregations because too labyrinthine a complexity is created.[16] We find that, in real urban aggregates, axial and convex scales tend to be increased in line with the scale of the growing object. For example, aggregates at the scale of a hamlet may be maximally two or three axial steps deep from their periphery. However, aggregates at the scale of towns are often also maximally two or three steps deep. The rules that govern growth change, it seems, with scale, in the manner of an allotropic process, and they do so in order to maintain certain parameters – for example that of depth – if not at a constant level, then at least at a level considerably less than would occur if the rule of growth did not change.

We call such rules 'globalising rules': they have the effect of maintaining the coherence of the growing global object from the point of view of what we might call the 'peripatetic subject' – a hypothetical individual moving around the system at ground level. Sometimes these globalising rules take quite precise forms. In the premodern City of London, for example, whichever City gate was entered, the 'centre' (that is, where Cheapside and the western radial routes meet) could be reached in three axial steps, provided only that at every point of choice the longest available line of sight was followed. One could teach an automaton to find the way to the City centre – perhaps because, for all intents and purposes, a stranger entering a town for the first time is a kind of spatial automaton.

More generally, we note, in the City of London, a tendency which we might call the 'conservation of shallowness'. Again and again, one finds that, for example, a back-alley route that, in plan, might appear tortuous, on the ground is not tortuous at all, because the axial organisation of space is far less complex than the convex organisation. Often there is a point where a line of sight goes back to the last major spatial event – say a major line, or a larger convex element – and another leads on to the next. Scaling axial lines upwards to pass through series of convex spaces seems to be the prime means by which this 'on the ground' intelligibility is created.

We also often find, in urban systems, a rather more complex principle which we might call the conservation of axial integration: the grid maintains the same degree of mean syntactic integration as a whole, as for its parts.[17] This 'conservation of integration' is found in cities with a high degree of integration, such as the City of London[18] and also in cities with very much lower degrees of integration, for example some of the cities of the Sahara and North Africa.[19] Arab towns are much less integrated than the City of London; but, like London, the mean integration of their parts, analysed independently, closely approximates the mean integration of the whole.

In many urban cultures, it seems, integration is a morphological constant, independent of the size of the system. The means by which this is achieved seems to be, once again, the globalising rules whereby axial and convex scale is increased in proportion to the scale of the system. The laws of urban growth that govern the constancy of integration within a culture appear to be more general than any tendency to specific degrees of integration. This suggests that the 'human scale' in space is not a metric absolute, but is relative both to culture and to the size of the urban system, and is an important means of maintaining intelligibility and functionality in urban space.

If urban space does tend to conserve certain cultural constants by the adjustment of scale, then may this also explain why we find linear, and transformable linear, relationships between axial integration and movement? The possibility is the more intriguing given the current preoccupation with nonlinear models: one feels one must justify even looking for linear relationships in the apparent chaos of human *peripeteia*.

Be that as it may, linearity from space pattern to functional outcome is, we suggest, not just a phenomenon which we discover as observers and investigators, but one which is built into the urban grid itself as an objective property. Urban grids, we suggest, evolve and grow in such a way as to ensure that natural movement is linearly predictable from spatial pattern, because the structuring – and therefore the predictability – of movement is the fundamental purpose of the grid; and the control of axial scale is the fundamental means by which the growing urban grid, within the laws of a particular culture, calibrates itself so as to generate and maintain a differentiated probabilistic encounter field with the interface properties specified by that culture.

We may develop this argument by looking more closely at fractal theory, and especially at its central concept of self-similarity: that is the propensity for systems to repeat, or nearly repeat, their forms at different hierarchical levels. To anyone familiar with the recent debate about architectural and urban space, the property of self-similarity will be familiar. The idea of a hierarchical nesting of similar or identical levels of spatial organisation, reflecting different levels of social organisation – individuals, families, groups of neighbours, neighbourhoods and so on – has been fundamental to how designers and theorists have sought to conceptualise the 'urban part' and how parts may be combined so as to form an urban whole. The theory that, at all spatial levels, humans formed groups which require a hierarchically separated environment (those based on 'human territoriality' were merely the most scientifically pretentious of such theories) has become the normative paradigm that has underpinned much twentieth-century design.

Natural movement studies, and syntactic studies of urban historical morphology, show how erroneous and how pernicious the assumption of hierarchical self-similarity has been in architecture and urban design. It has provided the theoretical support, now unstated, now explicit, for the creation of historically unprecedented urban grid, one preoccupied at each level with local separation and identity at the expense of global interrelationships, and as a result creating the fragmentary, over-localised and over-hierarchised enclaves that are the final urban product of the mid-twentieth-century design schools. In schools of architecture, indeed, the otiose notion of 'spatial hierarchy' has often been taken to be coterminous with order and structure in space.

But, however otiose, the notion of hierarchy is an attempt to describe a fundamental phenomenon of urbanism: that of the manifest existence of different scales of spatial organisation. How may this be understood without the simplifications of 'hierarchy'? Let us proceed empirically and consider the map of the City of London from the point of view of a 'peripatetic subject' – that is a hypothetical individual moving about the grid.

Imagine an individual walking eastwards up Cornhill from its western end. As the observer moves, he or she becomes aware of two levels of spatial organisation: one which relates him or her to the global scale of the grid and one which relates him or her to the much smaller-scale system of 'alleys and courts' of the block interiors. The peripatetic observer is given as many glimpses of the one as of the other. Although one scale is large and the other small, at both scales the observer sees a constant pattern in which axiality overcomes the tendency for space to become convexly localised.

As a result, both the intelligibility of the large scale, above the observer's level of immediate awareness, and the intelligibility of the smaller scale, potentially below the observer's level of awareness, are maintained by what appears to be a *principle of sufficient axiality*. Provided only that the observer continues to move, he or she will continue to experience movement as an interface between urban scales.

This two-way (from the moving subject) principle of sufficient axiality is not a metric invariant but a syntactic organising principle. Its effect is to maintain a kind of constancy between levels of space organisation that conserves intelligibility at the scale of the peripatetic subject. If space became too complex, above or below the level of the peripatetic subject, then intelligibility would be lost, and the most obvious route to this loss would be the loss of axiality.

As it seems logically necessary that the relation between spatial configuration and its functioning – that is, its predictability – must pass through the intelligibility of the configuration to the subject, then it also seems likely that the principle of sufficient axiality which conserves intelligibility is the means by which functional predictability is conserved as an objective property of the system. In other words, it is the axial organisation of urban space that ensures that people both understand it, and can intuitively predict its functional consequences.

Linearity, it might be speculated, arises from this, that is from the regular construction of the system of space in such a way as to conserve the relation of the peripatetic subject both to the small and to the large scales of space, and to keep both scales within his or her compass of understanding. The association between space configuration and natural movement thus becomes the prime determinant of urban spatial form, and the linearity of the relation between the two becomes an objective property of the urban system rather than simply a latter-day discovery by outside observers.

The ubiquity of the 'principle of sufficient axiality' requires us, I believe, to propose a redefinition of the human scale in a generally upwards direction. The 'human scale' means scaling spatial elements upwards to preserve the intelligibility and functionality of the system in accordance with the size of the system, and to keep both within the compass of the peripatetic observer. In this view, the human scale is about preserving the intelligibility of the inevitably large-scale systems within which we live, in spite of its complexity at the small scale and sheer metric size at the large. It is also a means of conserving the predictability of human movement from the grid, which is the essential purpose of the grid. By means of a constant principle – though within the confines of a given culture – axiality becomes the means by which linearity is, itself, rendered a constant. In other words, the issue of structure in space appears to be intimately bound up with that of scale, and vice versa.

One further speculation may be permissible. If attractors are found to have a logarithmic effect on movement, and if this may be taken to imply that the degree to which attractors appear on lines with a given degree of integration is also logarithmic, then it would surely follow that this would have implications for the scale of buildings, more specifically for their height. If the centre of a growing urban system maintains itself as the integration core of the system, then we should expect the height of its buildings to increase logarithmically in order to provide the attractor floor-space permitted by its degree of integration. If, on the other hand, the 'historic core' of the town is not maintained as the core of integration, then the increase in height would be much less. Both types of case would appear to exist.

The implications of natural movement for urban design are straightforward. If we wish to design for well-used urban space, then we must design with the knowledge that integration is a global variable, and movement in particular spaces is not determined, in the main, by the local properties of that space, but by its configurational relation to the larger urban system.

Methodological aids to achieve this in design are also straightforward, and are set out elsewhere.[20] To make sure that a new urban development works in terms of natural movement, one must first analyse the spatial and functional logic of the surrounding area by carrying out a computer analysis and observational study. Then, on the basis of the pattern revealed, candidate design proposals may be inserted into the computer model to analyse how they will create a pattern of integration within the development site, and how this will relate to that of the existing area. On this basis, reasonable predictions may be made of how the design will affect natural movement within and around the site.

In a sense, one might say that by emphasising natural movement, space syntax offers a normative idea of what constitutes good design and a successful outcome. However, experience suggests that there are many different ways to design a ghetto, but very few ways of designing an integrated system. Space syntax need only be invoked for the more difficult task. In this, however, it often offers no more than a powerful aid to the designer's intuition and intentions. It does not tell designers what to do. It helps them to understand what they are doing.

Notes

1. Respectively: G. K. Zipf, *Human behavior and the principle of least effort: An introduction to human ecology* (Cambridge, MA: Addison-Wellesley Press, 1949); J. Gibson, *The ecological approach to visual perception* (Boston, MA: Houghton Mifflin, 1979); D. Kahneman, *Thinking, fast and slow* (New York: Farrar, Straus and Giroux, 2011).
2. J. Appleton, 'Prospects and refuges re-visited', *Landscape Journal* 3 (1984): 91–103.
3. Paper originally presented at the *European Conference on the Representation and Management of Urban Change*, University Centre, University of Cambridge, UK, 28–29 September 1989.
4. For example, B. Pushkarev and J. M. Zupan, *Urban Space for Pedestrians* (Cambridge, MA: MIT Press, 1975).
5. B. Hillier, 'Against enclosure', in *Rehumanizing housing*, edited by N. Teymur, T. Markus and T. Wooley (London: Butterworth, 1988), pp. 63–85; B. Hillier, R. Burdett, J. Peponis and A. Penn, 'Creating life: Or, does architecture determine anything?', *Architecture & Behaviour* 3 (1987): 233–250.
6. That is, techniques which find regularities that form the basis for prediction.
7. The editors would like to point out the importance of this statement in the history of space syntax. The idea that space generates fields of encounter was present from the early days of space syntax. Here the finding that the configuration of the network *itself* shapes patterns of movement was an emergent discovery from research.
8. This may at first seem surprising because this representation is commonly used in transport modelling. However, in these applications the form of representation is relatively unimportant. The main factors in the success of these models are the assignment and calibration routines.
9. B. Hillier and J. Hanson, *The social logic of space* (Cambridge: Cambridge University Press, 1984); for a current discussion of this, see M. Krüger, 'On node and axial grid maps: Distance measures and related topics', in *European Conference on the Representation and Management of Urban Change* (Cambridge: 1989); working paper version available at https://discovery.ucl.ac.uk/id/eprint/1011/.
10. Hillier and Hanson, *The social logic of space*.
11. As shown in Hillier and Hanson, *The social logic of space*, and discussed by Krüger, 'On node and axial grid maps'.
12. J. Hanson, 'Order and structure in urban space: A morphological history of the City of London' (PhD diss., University of London, 1989), p. 549.
13. Hillier et al., 'Creating life'.
14. J. Peponis, E. Hadjinikolaou, C. Livieratos and D. A. Fatouros, 'The spatial core of urban culture', *Ekistics* 56 (1989): 43–55.
15. Hillier et al., 'Creating life'.
16. Hillier and Hanson, *The social logic of space*.
17. See Krüger, 'On node and axial grid maps'. The editors note that the original 'Natural movement' article lists, but doesn't cite, another Krüger paper: M. Krüger and B. Turkienicz, 'Synchrony in urban form', in *Third Seminar on Urban Design in Brazil* (Brasília: University of Brasília, 1988).
18. Hanson, 'Order and structure in urban space'.
19. A. Loumi, 'Spatial configuration of towns in North Africa' (PhD diss., University of London, 1988).
20. Hillier, 'Against enclosure'.

15 Specifically architectural theory (1993)

Introduction to 'Specifically architectural theory'

Ashraf M. Salama

A rich and unique representation of the discourse that argues for the essentiality of seeing architecture as both science and art, 'Specifically architectural theory', published in the *Harvard Architectural Review* in 1993, is a key text underpropped by empirical investigation and following several decades of experimentation in spatial analysis. It captures Hillier's positional blend of theory and action pertinent to the understanding of spatiality and cultural meaning as they relate to buildings and urban environments, and the society within which they exist.

This article establishes a distinct line of enquiry it moves from problematisation, to conceptualisation, to operationalisation in a lucid logical manner while utilising key examples of the work of Modernist including Adolf Loos and Le Corbusier. How does this manifest? Clearly, the argument is put forward as the case for a theory specific to architecture, one that emanates from within it, rather than being transposed from another field of enquiry or area of knowledge and then applied to it.

Hillier problematises the ontological nature of architecture and architectural theory and highlights the conceptual challenges of distinguishing architecture from 'the vernacular'. He contends that architecture is not only a cultural transmission or a reproduction of social knowledge; it is an outcome (product) and a judgement (process). In constructing the case for a *specifically architectural theory*, he conceptualises two notions of space – *the first* as a philosophical problem where he opines that space is never apathetic, rather a key constituent of the physical world that accommodates social and cultural processes, and *the second* as a pattern problem – while calling for the need to have a rational language that describes the forms and configurations resulting from these processes. Operationalising these conceptions into a case-based discussion on *abstraction* and *concretion*, an in-depth spatial analysis of domestic interiors authoritatively concludes that space does not lose its social dimension despite the evidence that it *moves outside the realm of specific codes of social knowledge*[1] and that the relation between spatial and social forms is interdependent and follows consistent patterns.

Considering the urban grid as a driver for architectural thinking, the analysis moves into articulations of the constituents of urban space and the challenge of identifying elements that can be subjected to configurational analyses, including aspects of continuity and irregularity.

Extending beyond the analysis which operationalises the theory, a dedicated section for designing with space syntax is included to provide the case of the King's Cross masterplan development. The analysis and the thought process involved in the development of the masterplan represent a manifesto for learning from the past without authorising its forms, shapes or patterns in their physical sense, emancipating contemporary principles from historical precedents. The period between the late-1960s and the mid-1990s witnessed significant developments in design and architectural theories. I argue that it is imperative to view this article from the lens of the intellectual context at the time it was written. The body of knowledge on design methods and theories that emerged from the 1960s onwards demonstrates the evolutionary nature of thinking about designing products and environments.[2] This was thinking that was originally considered through a scientific approach where design was depicted in an iterative-systematic, rationalistic manner and in terms of inputs and outputs. This iterative process was viewed as linear and simplistic, adopting the view of design as a well-defined problem. A reaction against this approach was developed to place emphasis on design as an ill-defined problem and on the characteristics of design contexts, contents and processes[3] and to address biases in information gathering and problem definition and structuring.[4] In this respect, design was seen as an interaction of basic ideas and modifying factors while users' input is regarded as essential to the process. This was advanced by a third approach that brought the competence of the designer into a better and more inclusive understanding of design. It adopted the view that design is a distinctive kind of thinking fundamental to human intellectual abilities, is a reflection-in-action process,[5] where the notions of tacit knowledge about design process and abductive reasoning for decision-making are employed.

'Specifically architectural theory' is well positioned within the preceding discourse, and, in essence, relates to the three approaches; it is premised on drawing clear boundaries between architecture and building and established the understanding of a thought process required to both instigate and apply a theory. By further articulating key concepts in design methodology, it expands (a) decision-making to judgement based on comparative thought and knowledge, (b) systematic intent to include theoretical intent, and (c) architectural competence to include 'universalistic' competence.

The concluding argument of the article furthers the discussion on design thinking as it relates to architecture and urban design by emphasising that architecture is both science which involves *abstraction* processes, and art which embroils *concretion* practices. Balancing, in Hillier's words, 'the creation of a theoretical realm' and 'the creation of an experienced reality' is fundamental to 'Specifically architectural theory'. Even though this article was published 30 years ago, the validity and currency of its content endure since architecture, as a discipline and as a profession, continues to struggle to define its own knowledge base beyond its 'vocational' one.

The outlook of 'Specifically architectural theory', I would deduce, is open to various interpretations on the prospects of knowledge in architecture and urban design. Prominently, the argument elucidates two types of knowledge. *Type A* is concerned with knowledge stemming from research that seeks to understand the future through a better understanding of what is there, the past: research that tests accepted ideas and revalidates conceptualisations. *Type B* is oriented to knowledge originating from research that seeks to probe new ideas and principles which could shape the future: research that develops new visions and verifies new hypotheses. Evidently, 'Specifically architectural theory' accommodates these two types while potentially enabling a spectrum of knowledge spaces in architecture and urban design, some of which are evident in many subsequent studies following similar lines of thought.

Specifically architectural theory

A partial account of the ascent from building as cultural transmission to architecture as theoretical concretion

Bill Hillier

Originally published: Hillier, B. 1993. 'Specifically architectural theory: A partial account of the ascent from building as cultural transmission to architecture as theoretical concretion'. *Harvard Architecture Review*, 9, 8–27.

Introduction: Architecture as theory

Theories are forms of knowing that summarise experience into abstract principles, and thus transform the meanings we assign to experience and the way we act on the world. Architects use theories in design, knowingly or unknowingly, not only because the creation of forms must reflect how the designer understands the world, but also because architecture, unlike everyday building, seeks as yet unknown forms, whose nature cannot, by definition, be predicted from experience.

But what are architectural theories like? Are they intellectual styles, like semiotics or deconstruction, brought into architecture from outside and interpreted within architecture? Or is there also some harder-edged sense in which architectural theories are specific to architecture, aiming to explain architectural phenomena as well as to guide design? Are architectural theories, in short, theories *applied* to architecture or are there also theories *of* architecture?

In this article, I argue that theories of architecture exist, and that they are to be found, not in the changing intellectual context of architecture as a bookish appendage to practice, but within the practice itself, guiding the answers to kinds of questions that arise at the point of design. Architecture, I will argue, is an intrinsically theoretical act. The key to architectural theory lies, I will suggest, not in the invocation of external abstractions, but in a proper understanding of the processes and products of architecture.

I will argue my case for specifically architectural theory using the problem of space as a specific instance. Architects and builders already use theory-like constructs in creating built space, and I will try to show that it is possible to develop a much fuller theory of space, one with some pretence to objectivity, capable of augmenting our intuitions in explaining and predicting forms, and also capable of refutation. I will argue that although such theories challenge architects with much more powerful and precise tools of analysis than they have had before, they lead not to constraint but to liberation.

Better theories of space mean more freedom for the designer because they bring the deep structures of architectural and urban space into the realm of rational debate and creative intuition.

In this article I will first try to distinguish architecture from building to show how theory is central to architectural practice. Then I will look at the issue of space, first as a philosophical problem, then as an aspect of buildings, and finally as an architectural phenomenon. I will then turn to the theory of space itself and suggest that space has its own internal laws, and that it is only when these are properly understood that space can be fully a part of architecture. Finally, I will draw some inferences on how this view of theory affects our view of architecture as science and as art.

Systematic intent of the architectural kind

First, how is architecture theoretical? Let us begin with some elementary semantics. If we try to unpack the ways we use the word *architecture*, it seems to refer both to an activity and a thing, that is, to the activity we call *design* and to buildings where we note evidence of this activity. Does this imply, as it seems to, that architecture is not really an objective property, but only a record of a certain kind of activity?

This is a difficult question, of a kind familiar to philosophers and aestheticians, who often ask whether words like *beautiful* refer to intrinsic properties of things, or are more akin to words such as *appropriate* which clearly do not refer to intrinsic properties of things, but to the judgements that we make by comparing things to other things.[6] Putting the question their way, we might ask whether architecture is actually a property of architectural objects, or a judgement that we make about objects, aware that they are the result of architectural activity.

Let us try to throw light on this by examining cases where deciding what is and is not architecture is particularly difficult, as when looking at the origins of architecture, or at where to draw the line between architecture and the vernacular. A colleague of mine, in reviewing the archaeological record for the origins of architecture, suggests that we see architecture in the evolution of buildings when we see evidence of 'systematic intent'.[7] By this, she means deliberate, abstract thought applied to construction, to space arrangement, or to visual organisation, either at the level of the building or the settlement.

This is an interesting and persuasive definition. But if we try to generalise it, we encounter problems. Suppose, for example, that we try to use it to distinguish architecture from the vernacular. It doesn't work, because clearly the vernacular is full of systematic intent. To make the matter even more difficult, the demarcation between architecture and the vernacular shifts with time, in that aspects of the architecture of one generation may reappear as the vernacular of another, and vice versa. These difficulties really do begin to make it look as though architecture is not at all an intrinsic property of things but a judgement that we make about things in the light of other knowledge.

However, if we look a little more closely at the vernacular, we find new possibilities. The outstanding work of Henry Glassie on vernacular housing adapts from Noam Chomsky a concept he calls 'architectural competence', which, he argues, underlies the architectural consistencies and variations by which we recognise a vernacular tradition. For Glassie, 'architectural competence' is a set of technological, geometrical and manipulative skills relating form to use, which constitute 'an account not of how a house is made, but of how a house is thought . . . set out like a program . . . a scheme, analogous to a grammar, that will consist of an outline of rule sets interrupted by prosy exegesis'.[8] Glassie's analogy with language is apposite. It suggests that the rule sets the vernacular designer uses are often tacit, taken for granted in the same way as the rule sets that govern the use of language. They are ideas we think *with*, rather than ideas we think *of*. The proposal that the evidence of systematic intent that we note in the vernacular might have its origins in some such rule sets seems a compelling one.

The implication of Glassie's idea is that 'architectural competence' provides a set of normative rules about how building should be done, so that a vernacular building reproduces a known and socially accepted pattern. The house built by a builder, sharing the culture of a community, comes out right because it draws on the normative rules that define the architectural competence of the community. Buildings become part of what Margaret Mead calls 'the transmission of culture by artefacts'.[9] Through distinctive ways of building, aspects of the social knowledge of a community are reproduced.

Now, whatever architecture is, it is clearly not just the transmission and reproduction of social knowledge through building, though it may include that. But this does suggest where the difference between architecture and building might lie. What we mean by architecture surely is not building by reference to *culturally bound competences*. What we mean, rather, is building by reference to a would-be *universalistic competence* based on general comparative knowledge of architectural forms and functions, and aimed (through understanding of principle derived from comparative knowledge) at innovation rather than cultural reduplication. The judgement we make, that a building is architecture, arises when the evidence of systematic intent is evidence of intellectual choice and decision-making exercised in a field of possibility that goes beyond cultural idiosyncrasy and into the realm of principled understanding. It is when we see in buildings evidence of this concern for the abstract comparability of forms that building is transcended and architecture is named.

We may then generalise and say that building is transcended and architecture is named where we note, as a property of buildings, some evidence, not only of systematic intent but also of theoretical intent, at least in embryonic form. In this sense, architecture transcends building in the same sense that science transcends the practical arts of making and doing. Architecture introduces into the making of buildings a more abstract concern for the realm of possibility created through theoretical concern. In this sense, architecture is theory applied to building.

The demarcation between the vernacular and architecture is then no longer problematic. The reproduction of existing forms, vernacular or otherwise, is not architecture, because it requires no exercise of abstract comparative thought. But by the same criterion, the exploitation of vernacular forms in the creation of new forms can be architecture, because it does involve such thought. Architecture is thus both a thing and a judgement. In the form of the thing, we detect evidence of systematic intent of the architectural kind. From the built evidence we can judge both that a building is intended to be architecture and, if we are so inclined, that it is architecture.

Space as a philosophical problem

Now space, I will argue, is one of the primary means by which the ascent from building as cultural transmission to architecture as theoretical intent is made, and is therefore one of the prime targets for architectural theory. This is to say that one aspect of the abstract comparability of forms in architecture centres on spatial form, which implies that space is, in some important sense, an objective property of buildings.

This is not obvious. Most of our common notions of space do not deal with space as an objective entity in itself but tie it in some way to human agency. For example, laymen tend to transcribe space as the use of space, or the perception of space, or concepts of space. Space as a thing in itself is harder to communicate. Common spatial concepts in architectural discourse are also similarly tied: personal space, human territoriality, spatial scale and so on.

Even in architectural concepts of space where space is unlinked from direct human agency, we still find that space is not independently described. The concept of spatial enclosure, for example, describes space by reference to the physical forms that define it. Without them, the space vanishes. This tendency finds its extreme expression in writers such as Roger Scruton, who think that the concept of space is a rather silly mistake made by rather pretentious architects who have failed to understand that space is not an entity at all, but merely the obverse side of the physical object, the vacancy left by the physical building. For Scruton, it is self-evident that space in a field and in a cathedral are the same thing except insofar as the interior built surfaces of the cathedral create the impression that the interior space has distinctive properties.[10] All talk about space is in error, because it can be reduced to talk about physical objects.

In fact, this is a quite bizarre view, since at a practical level, space is manifestly the saleable commodity in buildings. We build walls, but we sell and rent space. Are developers who advertise space at so much per square foot making a category mistake? Should they be offering to rent walls and roofs? Why then is Scruton embarrassed by the concept of space? Let me suggest that Scruton is making an educated error, one that he would not make if he had not been so deeply imbued with the Western philosophical tradition.[11]

The dominant view of space in Western culture has been one we might loosely call Galilean-Cartesian. By this I mean that the primary properties of objects are seen as their extension – length, breadth, width and so on – which are also their measurable properties. Extensions are the indubitably objective properties of things, independent of observers, unlike secondary properties such as 'green' or 'nice', which seem to depend in some way on interaction with observers. If extension is the primary property of objects, then it is natural to infer that it is also the primary property of the space within which objects sit. We can see this by the fact that when we take the object away from its space, its extension is still present as an attribute of space. Space is therefore generalised extension and, as such, the framework within which the primary properties of objects are defined.[12]

But once we see space as a general, abstract framework or background of extension, then we are doomed not to understand how it plays a role in human affairs, including architecture. Space is never simply the inert background of our material existence. It is a key aspect of how our social and cultural worlds are constituted in the physical world, and structured for us as objective realities. Space is not the neutral framework for social and cultural forms. It is built into those very forms.[13] It is because this is so that buildings can carry within their spatial forms the kinds of social knowing that Glassie notes.

Space as a pattern problem

But because space is built so pervasively into social and cultural life, we tend to take it for granted, to the point that its forms become invisible to us, and so much so that we have no rational language for the discussion of these forms. The only language is that of the forms themselves. If we wish to build a theory of space, then we must first learn this language – although in a sense we know it already – and learn to talk about space in a way that allows its form to become clear.

Let us begin by defining the problem clearly: as a *pattern* problem. Consider the two notional courtyard buildings of Figures 15.1a and 15.1b, showing in black the pattern of physical elements. Figures 15.2a and 15.2b show in black the corresponding pattern of spatial elements. The basic physical structures and

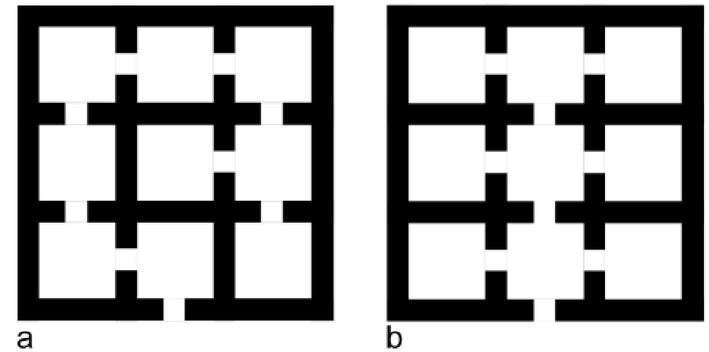


Figure 15.1 Arrangements of physical elements into structures.

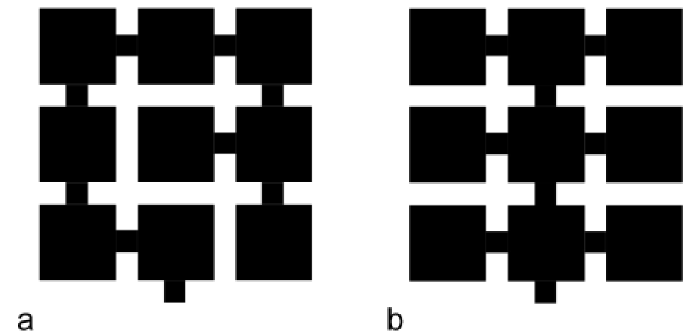


Figure 15.2 Arrangements of spatial elements into structures.

cell divisions of the two buildings are the same, and each has the same pattern of adjacencies between cells and the same number of internal and external openings. But the locations of cell entrances means that the spatial patterns are about as different as they could be from the point of view of the permeability of the layout. One is a near-perfect single sequence, with a minimal branch at the end. The other is branched everywhere about the strong central spaces.

The pattern of entrances would make relatively little difference to the building structurally or climatically, especially if we assumed a similar pattern of external fenestration and inserted windows in one building wherever the other had entrances onto the courtyard. But it would make a dramatic difference to how the layout would work as, say, a domestic interior. For example, it is very difficult for more than one person to use a single sequence of spaces. It offers little in the way of community or privacy, but much in the way of potential intrusion. The branched pattern, on the other hand, offers a more flexible set of potential relations between community and privacy, and many more resources against intrusion. These differences are inherent in the space patterns themselves in terms of the range of limitations and potentialities offered. They suggest the possibility that architectural

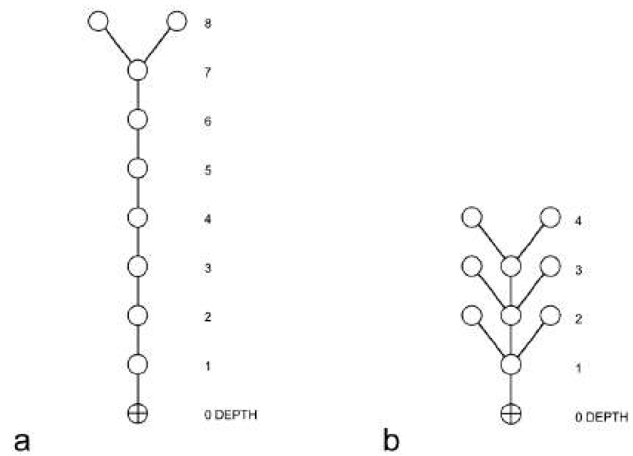


Figure 15.3 Justified graphs of permeability relations in Figures 15.1a and 15.1b.

space might be subject to limiting laws, not of a deterministic kind, but ones that set morphological bounds within which the relations between form and function in buildings are worked out.

We can capture the difference between the two spatial patterns by a useful device we call a justified graph (Figures 15.3a and 15.3b). In this, we imagine that we are in a space that we call the root or base of the graph, and represent as a circle with a cross. Then, representing spaces as circles, and relations of access as lines connecting them, we align immediately above the root all spaces that are directly connected to the root. Then above the first row, we align the spaces that connect directly to first-row spaces, and so on. The result is a picture of the *depth* of all spaces in a pattern from a particular point in it.

We can see that one is a deep tree form, and the other a shallow tree form. By tree, we mean that the patterns lack any rings of circulation. All trees, even those as different as the two in Figures 15.3a and 15.3b, share the characteristic that there is only one route from each space to any other – a property that is highly relevant to how building layouts function.

However, where rings are found, the justified graph makes them clear as depth properties (Figures 15.4a, 15.4b and 15.4c). Using justified graphs, then, we can begin to make visible two of the most fundamental properties of spatial configurations: how much depth they have from each space (how many other spaces must be passed through to get to others); and how each space relates to the pattern of circulation rings in the configuration (how it relates to the choices of route available).

More significantly, we can now take the crucial step in understanding spatial configuration as a product of culture. The key to spatial configuration in buildings and cities is that, within the same building or urban system, space has different configurational properties when looked at from different points of view. This can

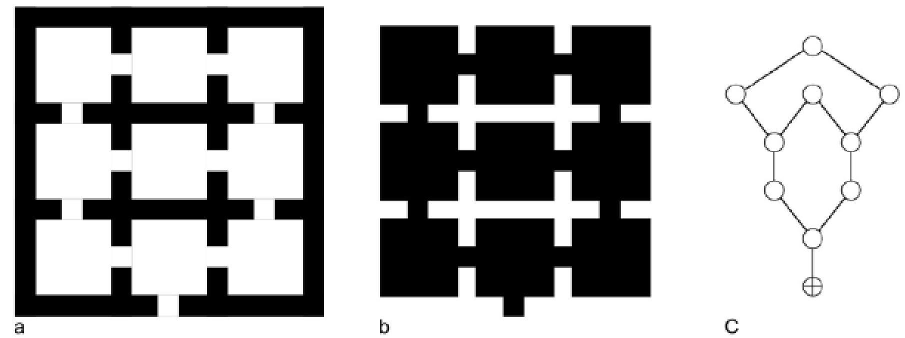


Figure 15.4 ‘Ring’ spatial configuration (a and b) and its justified graph representation (c).

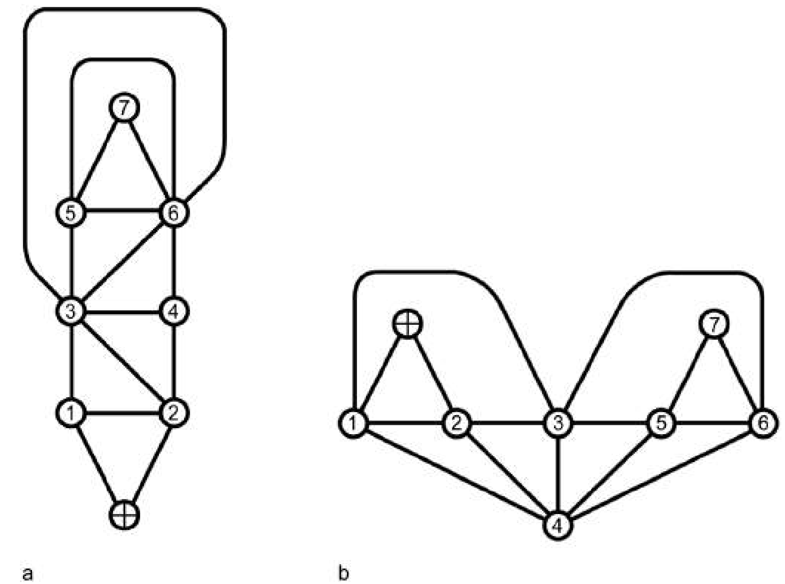


Figure 15.5 Graphs of the same spatial configuration justified from different points.

be shown by drawing justified graphs, because the differences have mainly to do with the way in which depth and rings are distributed in the spatial configuration when seen from different points of view (Figures 15.5a and 15.5b).

The depth and ring properties could hardly appear more different if they were different configurations. It is through the creation and distribution of such differences, by the arrangement of physical constructions, that space becomes such a powerful raw material for the transmission of culture through buildings and settlement forms, and also such a potent means of architectural discovery and creation.

Society in the form of the object

Let me now show how buildings can transmit cultural ideas through this aspect of spatial patterning. Figure 15.6 shows the ground floor plans of three French houses and their justified graphs drawn initially from the outside, treating it as a single space, then from three different internal spaces. Looking at the first graph (drawn from the outside), we see that in spite of the geometrical differences in the houses, there are strong similarities in the configurations. We see this most easily by concentrating on the space marked SC, or *salle commune*. In each case, the salle commune lies on all non-trivial rings (a trivial ring is one that links the same pair of spaces twice), links directly to an exterior space, and acts as a link between the living spaces and the spaces associated in that culture with the domestic work of women.

The salle commune also has a more fundamental property, one that arises from its relation to the spatial configuration of the house as a whole. If we count the number of spaces we must pass through to go from the salle commune to all other spaces, we find that it comes to a total that is less than for any other space – that

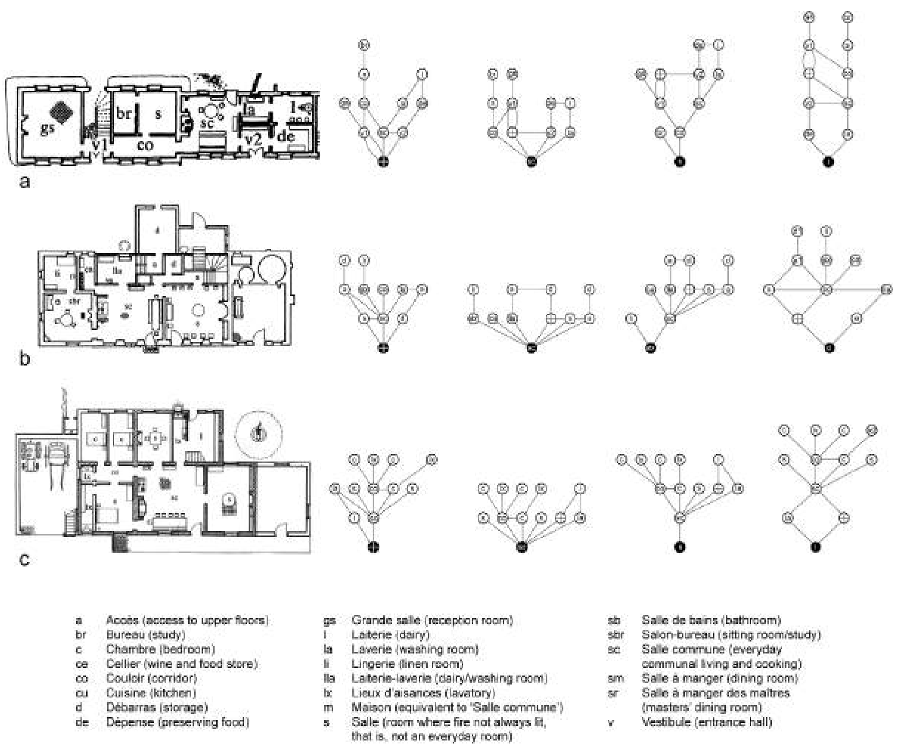


Figure 15.6 Plans of three French houses and their justified graph representations from different points.

is, it has less depth than any other space in the complex. The general form of this measure is called *integration*, and can be applied to any space in any configuration: the less depth from the complex as a whole, the more integrating the space.[14] This means that every space in the three complexes can be assigned an *integration value*. Other measures express how strong these differences are.[15]

Now once we have done this we can ask questions about the distribution of functions in the house. In the three French houses, for example, we find that there is a certain order of integration among the spaces where different functions are carried out, always with the salle commune as the most integrated. In other words, we can say with quantitative rigour that there is a common pattern to the way in which different functions are spatialised in the house. We call such common patterns *genotypes*, because they refer not to the surface appearances of forms but to deep structures underlying spatial configurations and their relation to living patterns.

These results flow from an analysis of space-to-space permeability. But what about the relation of visibility, which passes through spaces? Figures 15.7 and 15.8 show what we call the convex isovists (that is, all that can be seen from a space in which all points are mutually visible, in this case drawn to omit the corners of rooms in a consistent way) from the salle commune and another space labelled *salle*. In each case, the salle commune has a far more powerful visual field than the salle. These differences provide a basis for quantitative and statistical analysis

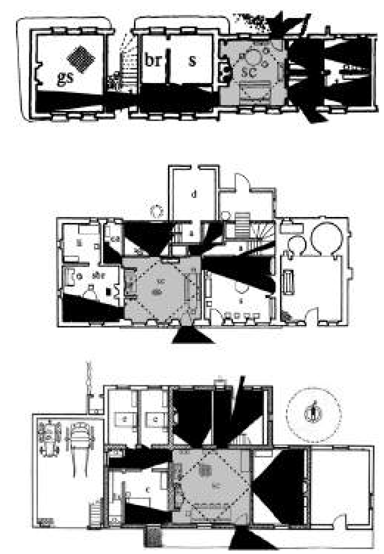


Figure 15.7 Diagrams of the houses from Figure 15.6 showing convex isovists from the *salle commune*.

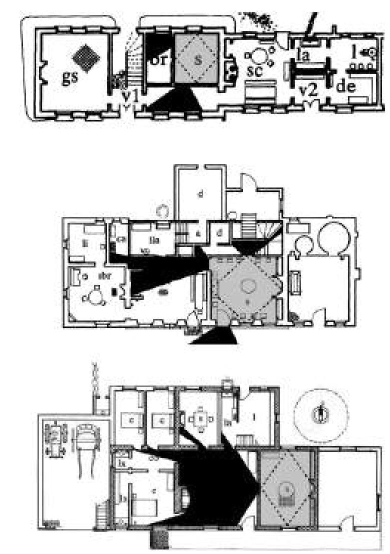


Figure 15.8 Diagrams of the houses from Figure 15.6 showing isovists from the *salle*.

and subsequent exploration of genotypical cultural patterns that lie embedded in the material and spatial objectivity of buildings.

This method allows us to retrieve what we might call Glassie properties from house plans, and to formalise the notion of cultural types. We have thus shown both how buildings can transmit social knowledge through their spatial form and how this can be retrieved by analysis. This is clearly useful knowledge for an architect to have. But it is not yet architecture, according to my definition, and certainly not a theory of architecture, even a partial one.

So how does this relate to the definition of architecture proposed earlier? Let me begin by referring to a study of selected houses by Adolf Loos and Le Corbusier by Dickon Irwin.[16] I cannot do justice to the subtlety and complexity of Irwin's argument in this brief text, but I would like to review some of his conclusions. Irwin's analysis of five houses by each of the two architects showed that although in each house there was configurational differentiation of functions, there was no consistent pattern within either architect's work. It was as though each recognised the principle that functions should be spatially differentiated, but regarded it as a matter of experiment and innovation, rather than as the reproduction of a culturally approved genotype.

However, Irwin was able to show that each architect had a distinctive spatial style, in that whatever each was doing with the functional pattern, distinctive spatial means were used to achieve the ends. For example, in the Loos houses, adding visibility relations to permeability relations increased the intelligibility of the space pattern,[17] whereas in the Le Corbusier houses it did not. Similarly, in the Loos houses, the geometry of the plan reinforced aspects of the spatial structure of the plan, in that major lines of spatial integration coincided with focuses of geometric order, whereas, again, in the Le Corbusier houses they did not. Some of these differences were captured by Irwin in diagrams he called *line isovists*, where he took the most integrated lines in the axial map of the house (see below) and drew all the space that could be seen from them. Figures 15.9 (Loos: Tristan Tzara House) and 15.10 (Le Corbusier: Villa Stein) show in order the isovists from the two most integrated lines from the point of view of the permeability pattern in each

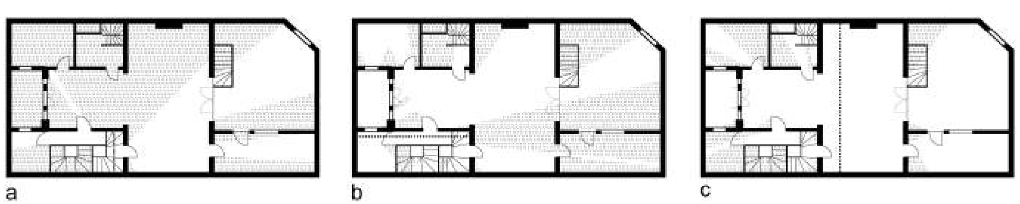


Figure 15.9 Adolf Loos, Tristan Tzara House. Fourth floor: (a) permeability isovist from the most integrated line; (b) permeability isovist from the second most integrated line; (c) isovist from the visibly most integrated line.

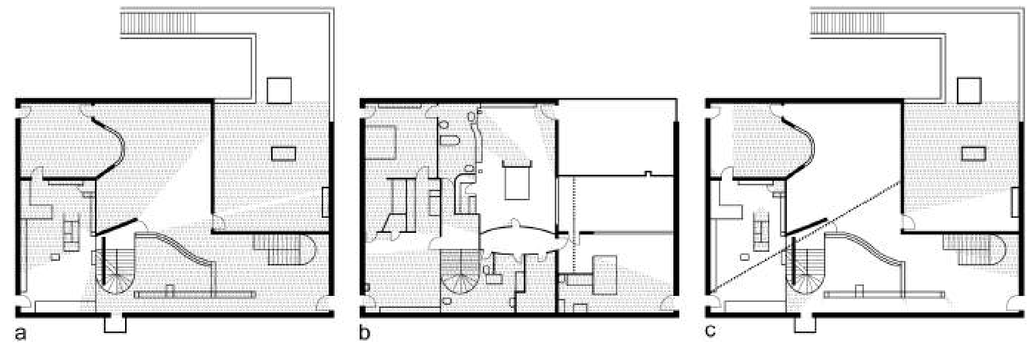


Figure 15.10 Le Corbusier, Villa Stein. Fourth floor: (a) permeability isovist from the most integrated line; (b) permeability isovist from the second most integrated line; (c) isovist from the visibly most integrated line.

house, followed by the visibility isovists of the two most visibly integrating lines. If we imagine each isovist as an episode in the spatial experience of moving through the houses, we can see that in Loos the isovists are very rich, but relatively uniform, whereas in Le Corbusier the isovists are more selective in the spatial relations they show from the line, but each episode is dramatically different from the others.

In this respect, Irwin argues, the two architects are adumbrating different fundamental – almost philosophical – programmes through architecture: Loos creates a house that is a novel expression of cultural habitability, Le Corbusier creates a less habitable, more idealised domain of rigorous abstraction. Neither Le Corbusier nor Loos is denying the social and cultural nature of the domestic interior. But each, by satisfying the need to give space cultural meaning through functional differentiation – first one way then another, but with a consistent spatial style – is giving priority not to the functional ends of building but to the architectural means of expressing those functional ends. The genotype of these houses lies, we might suggest, not in the functional ends, as in the vernacular cases, but in the way architectural means are used to express the ends. The means modify the ends by expressing them as part of a richer cultural realm.

This distinction between ends and means is, I believe, fundamental to the definition of architecture offered earlier. It suggests that we can make a useful distinction, in architecture as elsewhere, between the realm of social meaning and the realm of the aesthetic – in this case the spatial aesthetic. The cultural and functional differentiation of space is the social meaning, the spatial means is the spatial aesthetic. The former conveys a clear social intention, the latter an architectural experience that recontextualises the social intention. Meaning is the realm of constraint, the spatial aesthetic the realm of freedom. The spatial meaning of form expresses what architecture must be to fulfil its purpose as a social object, the spatial aesthetic expresses what it can be to fulfil its purpose as architecture.

But although space moves outside the realm of specific codes of social knowledge, it does not lose its social dimension. The relation between spatial and social forms is not contingent, but follows patterns that are so consistent we can hardly doubt that they have the nature of laws.[18] The spatial aesthetic carries social potential through these laws. The autonomy of architectural means thus finds itself in a realm, governed by general principles, with its freedom restricted not by the specific spatial demands of a culture but by the laws of space themselves.

These laws find one of their strongest expressions in urban space, where the social programming of space is much less closely defined than in building interiors. However, in looking for the operations and effects of these laws, we will find that certain attributes of urban space, believed by many to be aesthetic in origin, in fact arise from functional laws.[19] What these functional laws of space might be like is the theme of the next part of my argument.

The urban grid as an object of architectural thought

There are two factors that make the analysis of urban space especially difficult for configurational analysis. First, urban space is continuous. There is no obvious division into elements. Second, with obvious exceptions, urban space usually has a good deal of irregularity. Most towns and cities have deformed grids, with no obvious geometry. Both factors are aspects of the problem of representation: how do we define an element of urban space so that we can subject it to configurational analysis?

It will be useful to begin by looking at a familiar case and considering how we might think of urban grids as spatial patterns. Figure 15.11a shows, in black, the plan forms of all the open spaces and public squares in Rome, respecting orientation but not location. Figure 15.11b shows their location. Figure 15.11c shows the shapes, orientations and locations, and adds a further element: the full spatial shape visible from each square, its isovist. From this we see that some subsets of the isovists of the spaces form interconnected clumps with more or less continuous visibility and permeability, while others do not. These are pattern properties, arising from the interrelationships of many distinct entities.

How shall we analyse these Roman properties? The complexity of the situation is such that we must recruit the computer, and begin with some simple experiments. Figure 15.12 is a hypothetical arrangement of built urban blocks that create, by their disposition, an arbitrary deformed grid with a major, square-like space. Although deformed, it has a degree of continuity of space, which gives it an approximately urban look, unlike Figure 15.13, where the same blocks have been rearranged to create a pattern that is manifestly not urban. The difference is instructive. It tells us that the deformed grids we recognise as urban may have a good deal of internal order to them.

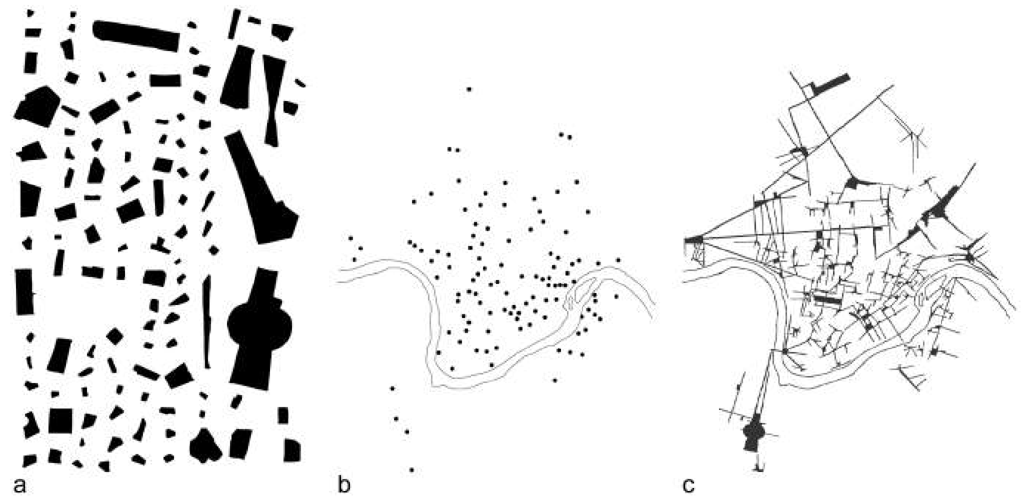


Figure 15.11 (a) Open spaces and public squares in Rome; (b) location of open spaces and public squares in Rome; (c) map of open spaces and public squares in Rome.

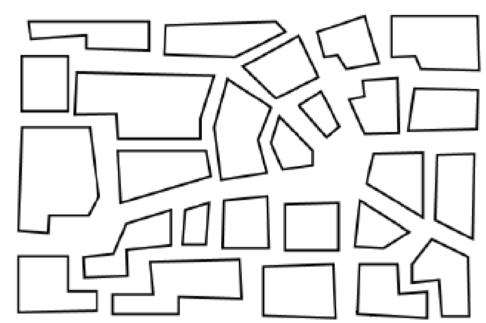


Figure 15.12 Hypothetical arrangement of built urban blocks.

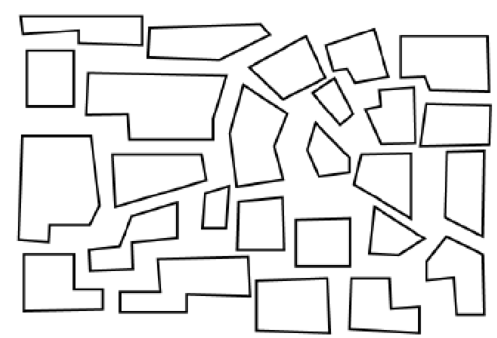


Figure 15.13 Hypothetical arrangement of built urban blocks.

Like other deformed grids, however, neither hypothetical figure has obvious spatial parts or elements. What, then, does it have that can be modelled? The answer is that, as we move about a deformed grid, it exhibits everywhere local properties that continually change. Just as the shapes of space that were experienced locally in the Le Corbusier or Loos houses changed as we moved through the house, so the shape of space we see as we move from point to point in an urban grid also changes.

The question is, how does it change? And can these changes be captured in a representation? The Roman case established several concepts that may be of use. For example, wherever we are in a deformed grid we are in some maximal, convex element of space defined by the surfaces of building blocks. The property of convexity means that any two points that can be seen from a point can also see each other. Figure 15.14 is a computer analysis of Figure 15.12 in which all such convex elements have been identified, allowing them to overlap as much as necessary, and then analysed and coded in terms of how *deep* each is from all the others: the darker, the less depth, the lighter, the more depth.[20] In other words, Figure 15.14 shows the distribution of *integration* (as defined above) in the convex representation of the deformed grid of Figure 15.12, with the darkest elements making up what we call the *integration core*.

Now let us carry out the same analysis on Figure 15.13. We find that the pattern of integration has changed completely, with the core moving to the edge and much weaker integration in the central areas (Figure 15.15). The integration core, one feels, has a much less powerful effect in creating an intelligible pattern to the space structure.

This property of intelligibility can in fact be demonstrated quite formally. If we go back to imagining that we are moving around in the spaces of these two configurations, then we can easily see that the field we see from points in Figure 15.15 will, on the whole, be a good deal less rich than the field in Figure 15.14. It will be a great deal harder to learn about the space structure as a whole because we get much poorer information from the localised parts. Few points in the pattern give much clue to the overall structure of the pattern, and even less to its distribution of

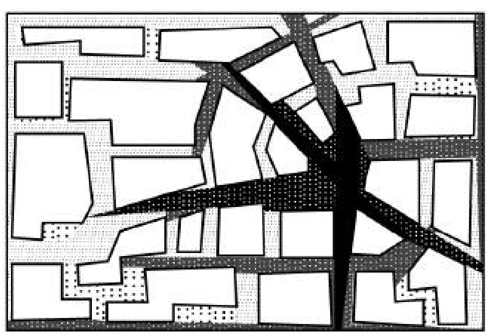


Figure 15.14 Computer analysis of the convex elements of Figure 15.12 showing distribution of integration.

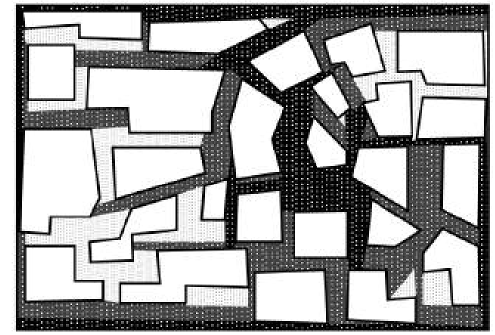


Figure 15.15 Computer analysis of the convex elements of Figure 15.13 showing distribution of integration.

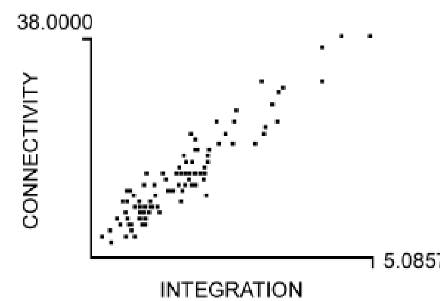


Figure 15.16 Intelligibility scattergram of Figure 15.12.

integration. In Figure 15.14, in contrast, we get a good deal of global information from local parts, and what we can see from points gives a good indication of how each space fits into the overall system.

This, in fact, reflects one of the most important pattern characteristics of deformed grid urban space. The information you get locally from the visual field you experience as you move around gives plenty of clues about how the overall spatial system is structured. In intelligible urban space, one might say, you get global information at the same time as you get local information about spaces, as we saw with the Roman squares.

Intelligibility can be quantified by a simple statistical trick. Figure 15.16 is a scattergram in which each point represents one of the overlapping convex elements of Figure 15.12. The number of convex elements each point overlaps with is indicated by its position on the vertical axis, and its degree of integration in the overall pattern on the horizontal axis. The more the points form a straight line from bottom left to top right, then the more connections an element has, which can be seen from each line. This is a reliable guide to its integration in the system as a whole, a property that cannot be seen from a line.

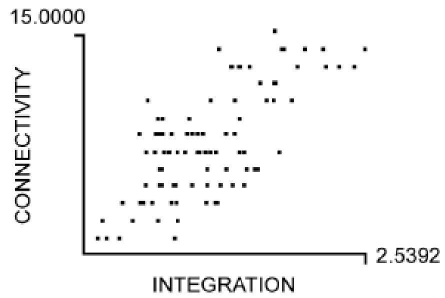


Figure 15.17 Intelligibility scattergram of Figure 15.13.

The tight scatter in Figure 15.16 shows that the first configuration has a high degree of intelligibility, which can be expressed as a number by taking the correlation coefficient of the scatter. A value of 1 would indicate a perfect straight line of points, and 0 a random scatter. If we now look at the scatter for the second configuration, shown in Figure 15.17, we can see that the scatter is much less tight, meaning that it has a lower value and therefore a lower degree of intelligibility. This expresses formally what intuition suggests: that the visual fields you see locally as you move around are a poorer guide to the system as a whole. The space structure is too labyrinthine. This analysis of the convex organisation of urban space is more than a formal game. It relates to important functional aspects of how space is used. For example, studies have shown that the choices that people make in selecting urban spaces for informal activities, such as eating, drinking, talking and sitting, reflect proximity or adjacency to areas with strong visual fields that are well integrated into the system as a whole.[21] Such spaces are ideally suited to what seems to be the favourite occupation of those using urban space informally: watching other people.

However, the most important and consistent functional effect of urban space follows from the configurational analysis of a different representation of its structure: one based on its one-dimensional, or axial, structure. We can again use the computer to explore the basics. The tangled skein of lines in Figure 15.18 represents the maximal linear visibility available within the open space structure of Figure 15.12, namely the set of all straight lines that are tangent to pairs of vertices of building blocks. That is, each line just passes by at least two such vertices, thus drawing a limit of a line of sight. Once the computer has found this set of lines, it can then subject them to integration analysis and code the results as before, with the darkest lines the most integrating, and showing the integration core of the pattern, and the lightest lines the most segregated (Figure 15.19).

Figures 15.18 and 15.19 thus represent different configurational views of the block arrangement in Figure 15.12. Each says: seen in terms of this type of local element, and analysed by that pattern parameter, the global structure of space looks like this. This is the essence of *space syntax* modelling. It is not a single technique

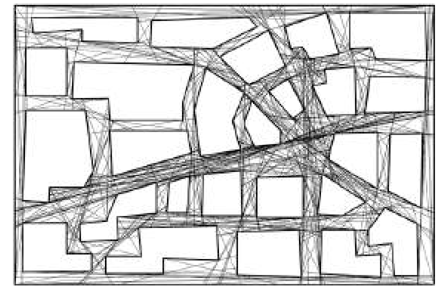


Figure 15.18 Diagram of maximal linear visibility available within the open space structure of Figure 15.12.

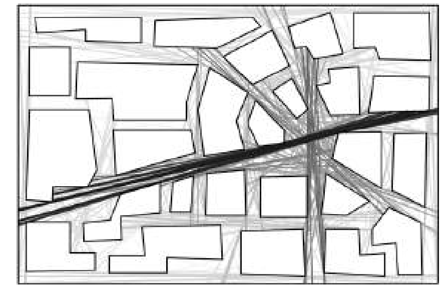


Figure 15.19 Computer analysis coding the diagram of maximal linear visibility shown in Figure 15.18.

but a set of techniques that allows two questions to be posed: how is the spatial system of interest to be represented as relatively localised elements?; and how are the interrelationships to be analysed to identify global patterns, so that we may understand the system's underlying structure?

Once we have this understanding of structure, we can begin to ask questions about function in a new way. Because syntactic analysis assigns, to each spatial element in a system, numbers that index its pattern relations, we can investigate the relation between these patterns and function simply by seeing how far the syntactic numbers assigned to spaces correlate with numbers describing aspects of function in those spaces: movement rates, informal use, rents, land uses, plot ratios and so on. We can thus pose questions about space and function in a new way. In the case of urban space, we can ask: what does function mean when space is universally public and more or less unrestricted? We will receive a resounding answer: urban space is about movement. Urban space creates a field of movement and thus co-presence and potential encounters among people.

We can show this by using again the scattergram technique. Figures 15.20–15.23 analyse the pedestrian and vehicular movement for the Barnsbury area of North London. The high degree of correlation in Figure 15.22 (the correlation coefficient is 0.85 on a scale of 0 to 1) shows that the number of people passing along each line is largely a function of the spatial pattern itself. The same is true of vehicular movement, whose scattergram shows a correlation of 0.81 in spite of the existence of a number of one-way systems (Figure 15.23). The fundamental result is that the pattern of movement in an urban system is determined, in the main, by the spatial configuration itself, and in particular by the distribution of spatial integration in the axial map of the system.[22]

These results are quite fundamental to our understanding of urban space, since they show that it is the *architecture* of the urban grid itself that is chiefly

Figure 15.20 Map of the area of Barnsbury in North London circa 1970 showing Barnsbury area of Islington.

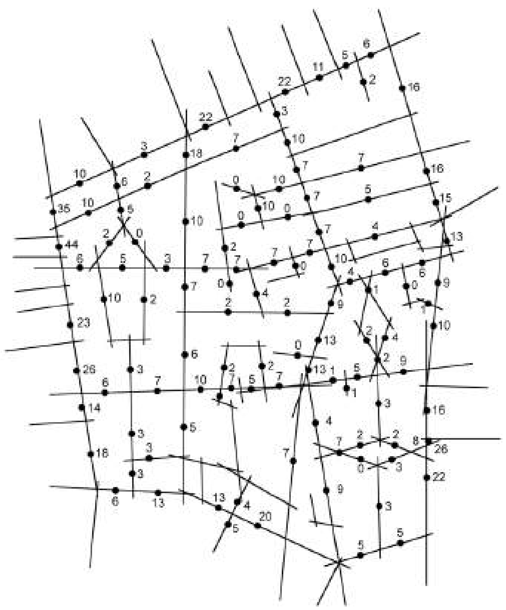


Figure 15.21 Axial map of the Barnsbury area. Numbers represent pedestrian movement rates per five minutes.

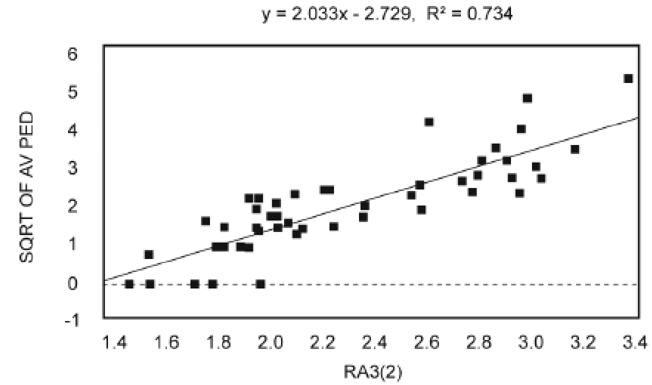


Figure 15.22 Scattergram plotting pedestrian movement against degree of integration. Degree of correlation is 0.85.

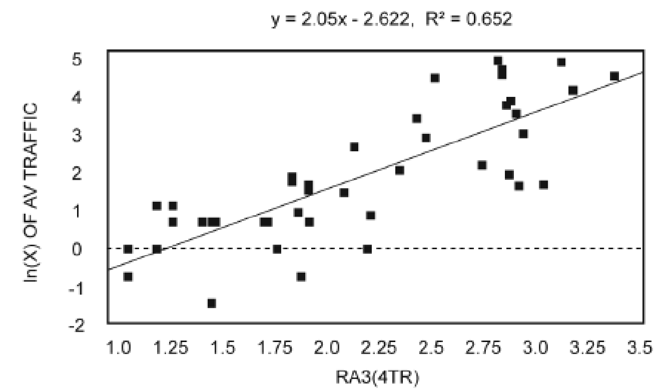


Figure 15.23 Scattergram plotting vehicular movement against degree of integration. Degree of correlation is 0.81.

responsible for the pattern of movement, not the positioning of attractors and magnets, as has commonly been believed. These results have been repeated so often that we have little doubt that they are something like a law. However, the law does not simply say that movement in a grid is a function of the distribution of spatial integration in the grid. The relationship, it turns out, is subject to the degree to which the grid has the property of intelligibility, as defined earlier. If you make urban space unintelligible, then you are also likely to make it unpredictable. We call this the theory of natural movement. Natural movement is the proportion of movement determined by the architecture of the grid itself. Where there is no natural movement, then most space will be empty for most of the time, leading almost inevitably to one aspect, or another, of urban malfunction. This is the reason why we must once again learn to make the urban grid an object of architectural thought.

Because natural movement is fundamental, it is also reasonable to suppose that it accounts, to a great extent, for the way in which urban grids evolve. It is likely that, over time, a dynamic relation develops between the evolving urban grid, its natural movement patterns and the developing pattern of land use. Certain types of use, for example retailing, survive best in locations that are accessible and have through-movement – that is, in locations that have both the spatial properties and functional effects of integration. The result is that, over time, urban grids evolve not only to optimise patterns of mutual accessibility, but also to optimise the usefulness of the by-product of movement from place to place – that is, the spaces that must be passed through on journeys from all origins to all destinations. Through this mechanism, spaces that are accessible for to-movement also become those with strong through-movement, and these spaces then become the busy foci of urban life. We call this the theory of the movement economy. If the theory is correct, it means that the architecture of the urban grid is of far greater significance in urban evolution than has been allowed in planning theory, and provides further reason for bringing back the urban grid as an object of architectural analysis and creativity.[23]

Designing with space syntax: The King's Cross masterplan

What does all this imply for design? Let us proceed through a real case study: the design by Norman Foster Associates of the masterplan for the King's Cross development in inner London for the London Regeneration Consortium. King's Cross is currently the biggest urban development project in Europe. Our published research using space syntax to predict pedestrian movement patterns, and the involvement of space syntax in public inquiries on major urban redevelopment schemes, had alerted first the community groups, then the planners and developers to the potential of using space syntax to help solve the fundamental problem of the King's Cross site: how to design the development in such a way that it continued and joined the urban structures of Islington to the east of the site and Camden to the west. Natural pedestrian movement to and through the site was seen as essential to this aim. Foster Associates, backed by the developers, asked us to make a study, and work with the design team in trying to build these relationships into the masterplan.

The first step was to study the spatial structure, space use and movement patterns in the existing contextual area. This study is documented in Figures 15.24–15.30. From a design point of view, the key product of the study is a spatial model of the contextual areas of the site, verified by its power to 'post-dict' the existing pattern of movement around the site. This allows us to add design proposals to the model, and to reanalyse in order to see how each proposal is likely to work within, and affect, the urban context.

We can therefore begin to explore intuitions as to what kind of masterplan will most successfully adapt the existing structure of the area and create the levels of

Figure 15.24 Block plan of part of the 1.5 km sq area of North London surrounding the site studied for the King's Cross project. Because natural movement is most strongly affected by the large-scale spatial pattern of an area, the contextual study area for a new development must not only be large enough to cover the likely pedestrian catchment area of the development, but it must also cover the catchment area of the catchment area, in order to ensure accuracy in the investigation of the catchment area. These two levels are modelled independently as the 'small area' – usually about 5–6 km sq – and the 'large area' – up to 20 km sq – to check that both levels are giving the same story (large cities have no natural internal boundaries) and also to ensure that any 'edge effect' – that is, distortion in the analysis resulting from the fact that some parts are close to the edge of the area modelled – is kept to the edge of the large area and does not affect the immediate contextual area of the site.

Figure 15.25 Black-on-white representation of Figure 15.24, showing the space of public access in black. The small-scale complexes around the site are housing estates.

natural pedestrian movement requested by the designers. This will depend on the achievement of two spatial objectives: bringing adequate levels of integration into the site in a way that reflects and adapts the existing natural movement patterns in the area, and maintaining, or if possible improving, the grid's intelligibility.

There is, however, a technical problem with the formal definition of intelligibility. Because intelligibility measures the degree of agreement between the local and global properties of space, a small system is, other things being equal, more likely to be intelligible than a larger one. We can overcome this by bringing in a database of established London areas of different sizes to compare with King's Cross

Figure 15.26 Axial map of Figure 15.25, showing the least set of longest straight lines of sight and direct access that pass through all the public space shown in the previous figure, omitting housing estates for simplicity. The axial map is the model of the area analysed by the computer to establish its underlying patterns. Of these, the most important is integration, that is, the mean depth of each line from all others in the system. The integration value of a line (or an intersection) is in effect the number of other lines that must be used to go from that line to all others in the system.

Figure 15.27 Axial map of Figure 15.26 analysed in terms of its distribution of integration, with the most integrating lines shown darkest, graded towards the least integrating shown lightest. Since integration predicts natural movement, one can think of the blackness of the line as predicting the amount of movement down that line.

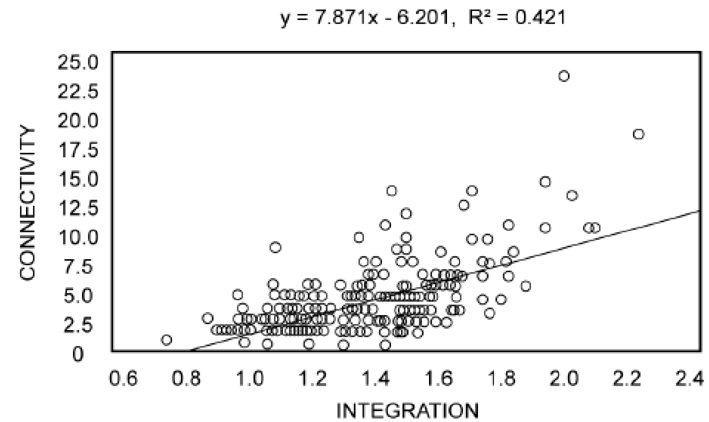


Figure 15.28 Intelligibility scattergram of the area as it stands, in which each line in Figure 15.27 is represented as a dot and located on the horizontal axis according to its degree of integration (a global property that cannot be seen from the line) and on the vertical axis according to its degree of connectivity (a local property that can be seen from the line). The index of intelligibility is the square root of the number at the top right. The scattergram shows a rather poor level of intelligibility, partly due to the hole in the system formed by the King's Cross railway development site.

as it is now and as it will be when it is developed (Figure 15.31). Figure 15.31 shows that not only is the King's Cross area less intelligible than London areas in general, but also the small area is relatively less intelligible than the large. This is probably because the 'urban hole' created by the King's Cross site has a stronger impact on the area immediately around it than it does on the larger surrounding area.

We can now use the area axial map as a basis for design simulation and experimentation. In fact, what we will be doing in this text is conducting a number of experiments that explore the limits of possibility for the site.[24] Let us first suppose that we impose a regular grid on the site, so that it has a local appearance of being ordered but makes no attempt to take advantage of the existing, rather disorderly pattern of integration in the area (Figure 15.32a). In spite of its high degree of internal connection, the scheme acts as a substantial lump of relatively segregated spaces, rather like one of the local housing estates shown in Figure 15.25, which freezes out virtually all natural movement and creates a quite unnerving sense of emptiness.

In other words, the grid scheme completely fails to integrate itself into the area or to contribute to the overall integration of the area. The effect on intelligibility is no better. We can see this by plotting the intelligibility scattergram and using the space syntax software to locate the spaces of the scheme on the scatter. Figure 15.32b shows that the scheme's spaces form a lump (within the box) well

Figure 15.29 Map showing 239 line segments in 10 areas around the site (including three housing estates adjacent to the site that are not shown on the axial map) where pedestrian space use and movement was observed using a simple moving observer technique, and distinguishing only between moving and static pedestrians, and between men, women and children.

off the line of intelligibility, occupying the segregated and rather poorly connected part of the scatter. We conclude not only that the scheme is far too segregated to achieve good levels of natural movement, but that its spaces are insufficiently integrated for their degree of connection and therefore worsen the intelligibility of the area as a whole.

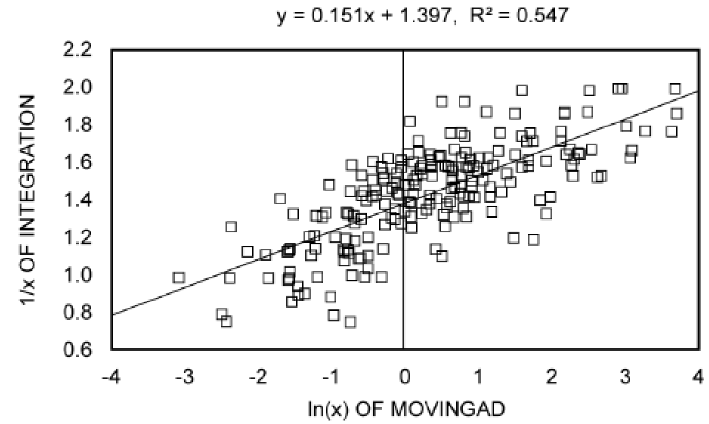


Figure 15.30 Scattergram showing each observed line in Figure 15.27 as a dot located on the vertical axis according to its degree of integration and on the horizontal axis according to the observed rate of pedestrian movement along the line. The correlation coefficient is .74 on a scale between 0 (no relation between the degree of integration and pedestrian movement) and 1 (a perfect relation between the two), showing how strongly the pattern of the urban grid itself influences the pattern of pedestrian movement.

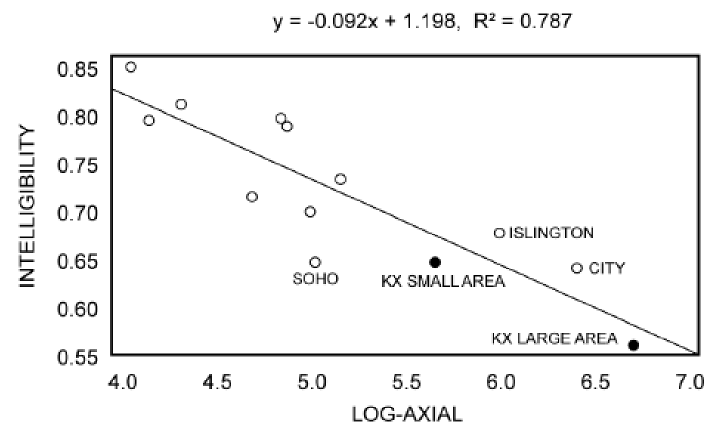


Figure 15.31 Intelligibility index plotting intelligibility against the size of system for selected areas of London. The regression line represents the established degree of intelligibility for an area of a given size. Of the four areas at the bottom right, the farthest right, and below the line, is the King's Cross area shown in Figure 15.24; next and above the line is the City; next, also above the line, is Islington; next following that, and below the line, is a smaller area around the King's Cross site.

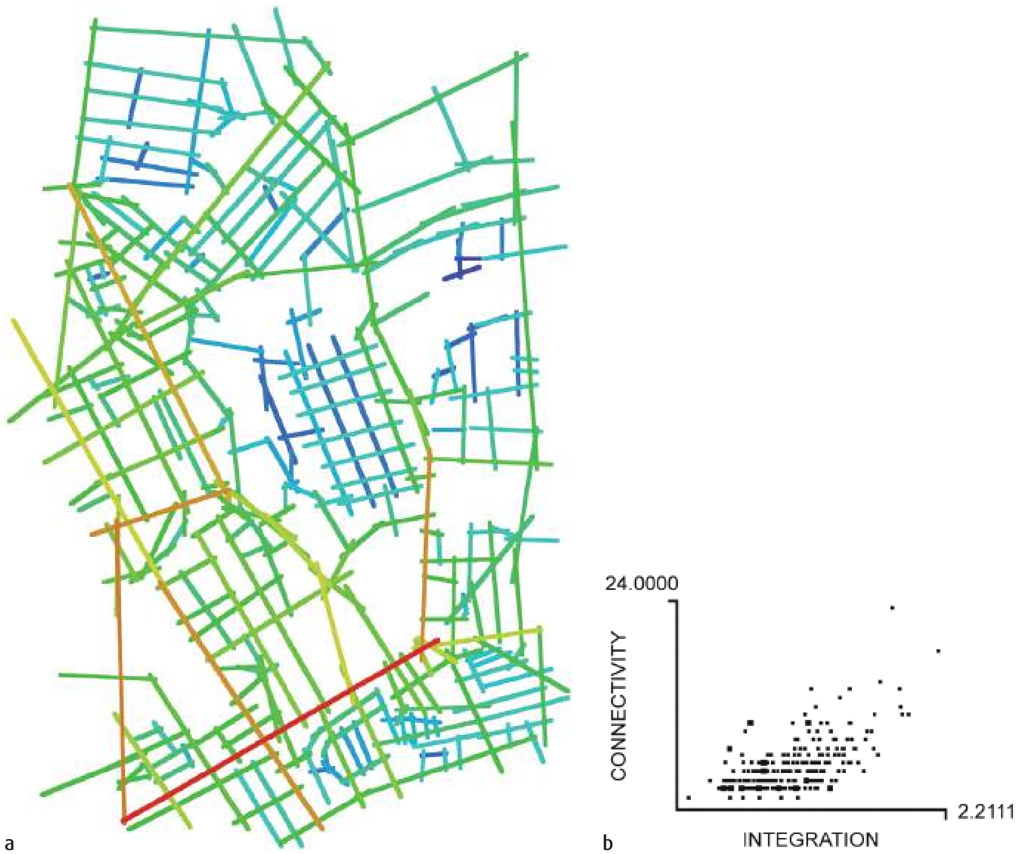


Figure 15.32 (a) Grid scheme analysed in its urban context to show how it affects the pattern of integration. Colours show integration in a range from red to blue, high to low values. (b) Intelligibility scattergram of grid scheme shown in Figure 15.32a.

Still in the spirit of experimentation, let us now try the opposite and simply extend integrating lines in the area into the site, and then complete the grid with minor lines more or less at will. This means that instead of imposing a new conceptual order onto the site regardless of the area, we are now using the area to determine the structure of space on the site. We must stress that this design is impossible, since it would require, among other things, a ground-level train crossing at the exit of St Pancras station!

In spite of its unrealism, the experiment is instructive. Figure 15.33 shows the integration structure that would result from such a scheme. In effect, it shows that certain lines extended across the site would become the most integrating lines in the whole area, stronger even than Euston Road – the major east-west trunk [main] road passing south of the site, which would be the major integrator with respect to

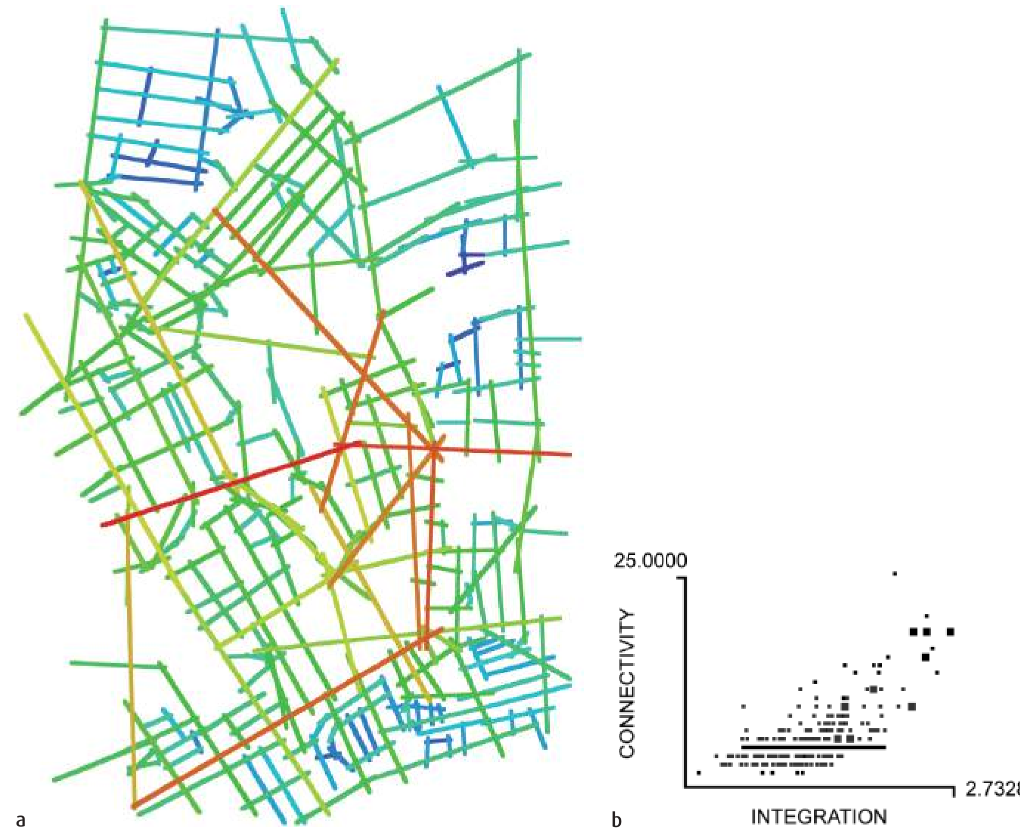


Figure 15.33 (a) Super-integrated scheme analysed in its urban context to show how it affects the pattern of integration. Colours show integration in a range from red to blue, high to low values. (b) Intelligibility scattergram of super-integrated scheme shown in Figure 15.33a.

an area expanded to the south and west. We also find a substantially greater range of integration values in the development, in contrast to the much greater uniformity of the grid scheme.

This is a 'good' urban property. Mixing adjacent integration values means a mix of busy and quiet spaces in close association with each other, with the kind of rapid transitions in urban character that are very typical of London. Figure 15.33b shows a much-improved intelligibility scattergram, with the lines of the scheme picked out in bold, showing that they not only improve the intelligibility of the area, but make a linear – and therefore intelligible – scatter themselves.

Now if we plot these two hypothetical solutions on the London Intelligibility Index (Figure 15.34), we find that, whereas the grid leaves the area as poor in intelligibility as it was, still lying in more or less the same position below the regression line, the 'super-integrated' scheme moves well above the regression line,

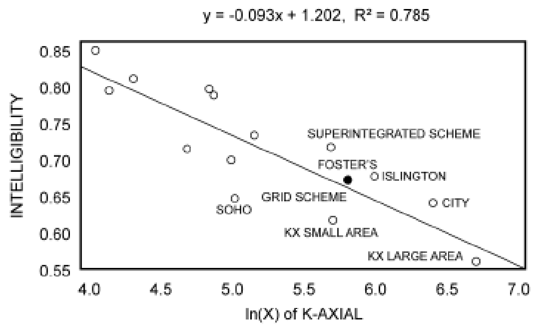


Figure 15.34 Intelligibility Index comparing grid scheme, super-integrated scheme and final Foster Associates scheme.

and would even be above the line formed by Islington and the City. We might even conclude that we have overdone things, and have created too strong a focus on the site for the mixture of commercial and residential development that is envisaged.

The final Foster Associates masterplan, working as it does within the concept of a central park to bring democratic uses into the heart of the site, is a much more subtle and complex design than either of these crude illustrative experiments. In developing the design, a protracted process of design conjecture and constructive evaluation through space syntax modelling took place, much of it around the drawing board. The final masterplan (Figures 15.35–15.37) draws integration in to, and through, the different parts of the site to a degree that matches the intended land-use mix, which ranges from urban offices and shopping areas, where levels of natural movement need to be high, to quieter residential streets, where levels of movement will be lower.

The intelligibility scattergram shown in Figure 15.37, again with the masterplan lines picked out in bold, shows that the scheme improves the intelligibility of the area, and also has high internal intelligibility, seen in the linear scatter of the masterplan lines. But the scheme also has more continuous variation from integrated to segregated than the super-integrated scheme, with its markedly more lumpy scatter. This indicates that the local variation in the syntactic quality of spaces, arising from mixing integrated and segregated lines in close proximity, is also better achieved. This is confirmed by the overall intelligibility index, which shows that the scheme falls very slightly above the regression line, meaning that it continues the established level of intelligibility in the London grid (Figure 15.35).

Thus, the design team may not only use space syntax to experiment with design in a functionally intelligent way, but they may also use the system to bring to bear on the design task both detailed contextual knowledge and a relevant database of precedent. We think of this as a prototype *graphical knowledge interface* for designers – meaning a graphically manipulable representation that also accesses contextual knowledge and precedent databases relevant to both the spatial structure and functional outcomes of designs. The experience of using space syntax on King's

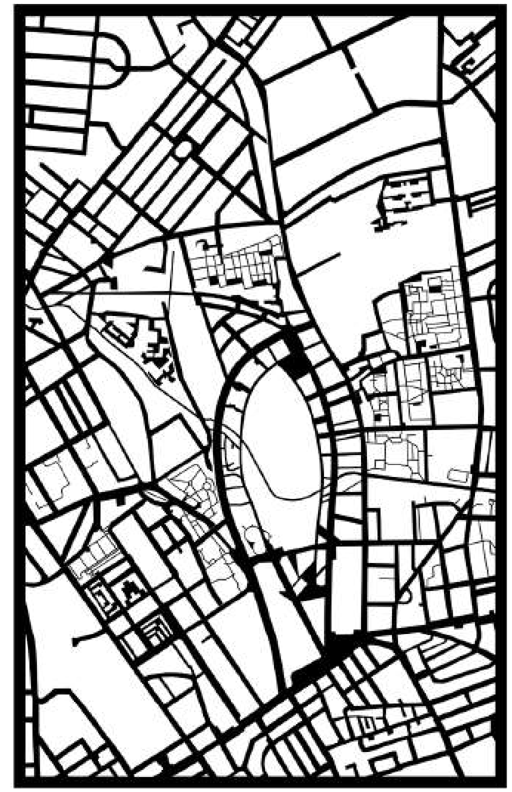


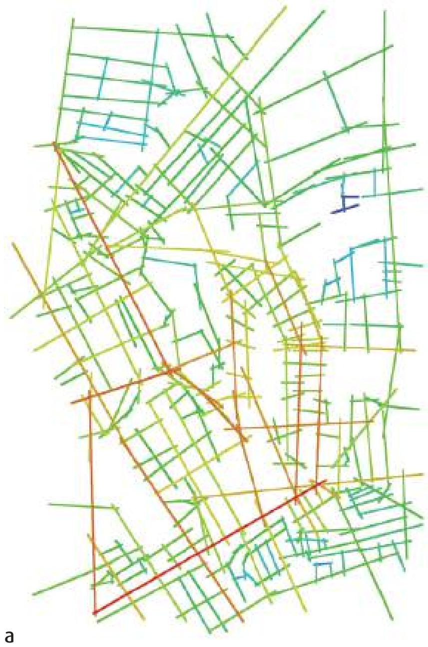
Figure 15.35 Final scheme for Kings Cross masterplan shown in urban context.

Cross and other urban masterplans has convinced us that what designers need from research is theoretical knowledge, coupled to techniques, not information nor data nor constraint. Furthermore, with theory and technique, much more of the living complexity of urban patterns can be brought within the scope of architectural intuition and architectural intent, without subjecting them to the geometrical and hierarchical simplification that have become the commonplaces of urban design.

Space syntax as a partial theory of architecture

If these are the implications for design, what then are the implications for specifically architectural knowledge and specifically architectural theory? There are two issues here. One concerns the forms that architectural theories and architectural knowledge take, the second how we conceptualise the relation of knowledge to design.

Regarding the first issues, it seems to me vital that space syntax theories are expressed in architectural form. By this I mean that theoretical knowledge is



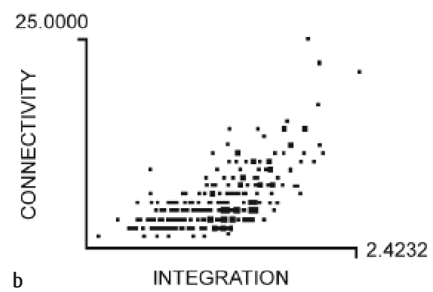


Figure 15.36 (a) Final scheme for King's Cross masterplan analysed in its urban context to show how it affects the pattern of integration. Colours show integration in a range from red to blue, high to low values. (b) Intelligibility scattergram of final scheme shown in Figure 15.36a.

brought to bear on the design through a form of representation that is directly architectural, not only in the sense that it actually copies, and allows manipulation of, aspects of architectural forms, but also in the sense that it carries within itself, through theory, knowledge of functional consequences. Syntactic representations are *theoretical descriptions*,[25] in that like buildings themselves, they are spatial forms with functional implications. Syntactic theories are architectural not only in the sense that they are *about* architecture, but also in the sense that they are in the language *of* architecture.

As for the second issue, the relation of knowledge to design, let us review the ascent from the vernacular to architecture. What we have seen is a series of levels at which we find theory-like entities in architecture. There is the level of the abstract social knowledge built into the 'architectural competences' that underlie the vernacular. Next is the level of the abstract typological comparison of forms. Then there is the level of general theoretical propositions, such as the theory of natural movement.

What is clear from the design application is that the most useful form of abstraction for design is the third level, that of general theoretical propositions. It

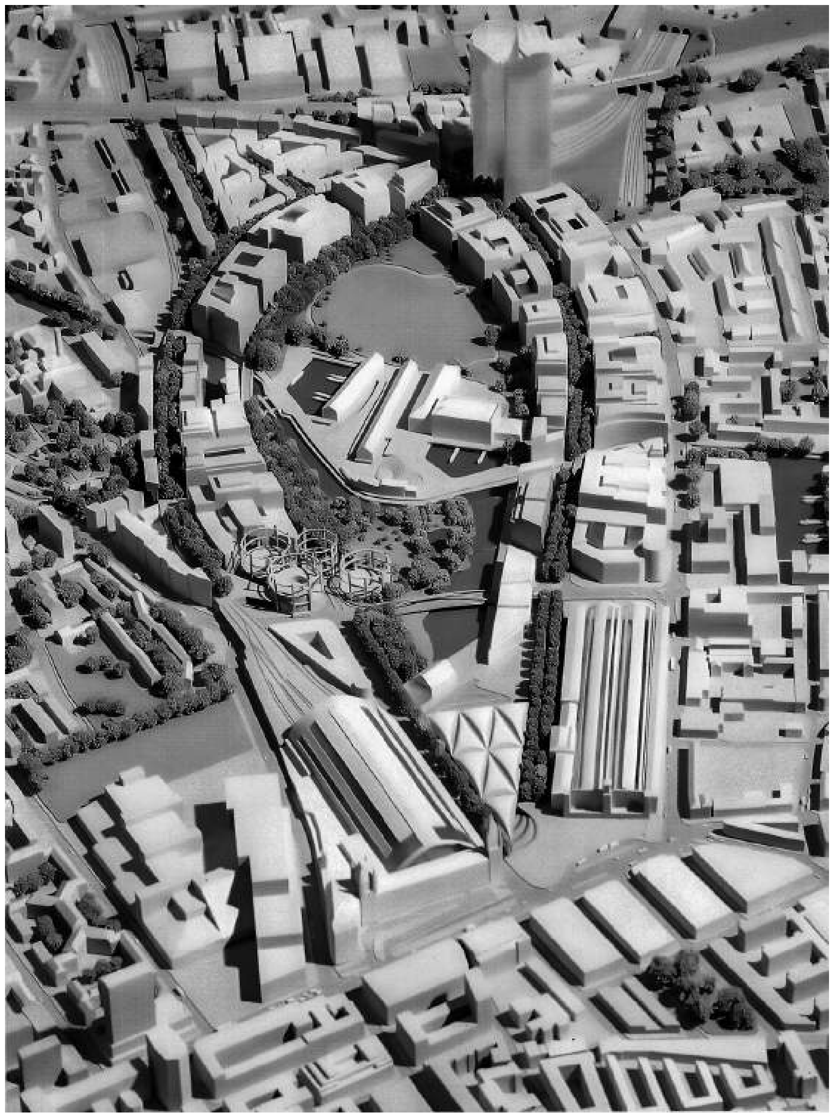


Figure 15.37 Norman Foster Associates, King's Cross masterplan, photograph of the model. Credit: Ruth Conroy Dalton.

is only at this level that strategic design thinking takes place – for example, about how a socially desirable, functional outcome of design, such as the integration of a new neighbourhood, might be achieved. It is also clear that theories generated by space syntax do not enter a theoretical void, but challenge theoretical ideas that already hold the field. Thus, in different ways, space syntax theory challenges notions like territoriality, defensible space, spatial enclosure, legibility through landmarks, geometrical theories, and a whole panoply of ideas about cities that currently play the role of theory in urban design.

Space syntax even challenges how design questions are defined. For example, a current topic of debate in Europe is the degree to which future development should be based on the past. In historic centres, for example, there is a widespread fear of doing anything except keeping the old street system, in spite of the obvious criticism that the street system to be conserved was created by a dynamic process of growth and change as each generation modified what it inherited to meet its needs and passed it on to the next generation. Conservation leads to the paradox that to freeze this process at one point in time, to conserve a specific form, would be antihistorical, since it would conserve a product but violate the process that gave rise to the product.

Space syntax redefines this question by making the issue one of genotype rather than phenotype. We can now ask not whether we should preserve specific forms, but whether we should preserve the underlying principles of specific forms in the light of present needs, or adapt them in the direction of a new genotype. History is replete with examples of both, in that, as cities evolve they can change their genotypes as well as their phenotypes.[26] What history does not offer is a precedent for the current fashion of phenotype conservation.

By showing how we can understand urban space genotypes, space syntax does allow the genuine continuation of a historical tradition without necessarily copying its surface forms. It does so by suggesting what is essential and what is inessential in the structures of the past. The King's Cross masterplan is a genotypical continuation of the logic of London. Yet it resembles no known part of London. It extends the deep structure of the existing grid, not its surface structure.

The tighter forms of reasoning, permitted by space syntax, thus have a liberating effect, precisely because they allow us to oppose the superstitious following of an established vernacular with abstract reasoning about forms and their functional consequences within an evolving structure. The intervention of theory, in effect, permits us to set the argument about history at the level of the evolutionary processes that generate the architecture of the city, rather than at the level of its specific products. The theoretical ascent from the vernacular, as social reproduction, to architecture, as the knowledgeable exploration of form through theory – even partial theory – is thus also an ascent from social constraint to liberation. Design can seek its goals not within the stultifying constraints of particular forms of social knowledge (which nevertheless can be and must be understood) but within the limits posed by the laws of architectural and urban space, and their realisation within a particular context.

Architecture as science

This redefinition of theory in terms of liberation is not obvious. At first sight, theories seem to be abstract schemes of thought that constrain rather than liberate. They appear to fix the mind in a certain way of looking rather than opening up new possibilities. However, this is to misunderstand the nature of theories, and their potential in architecture.

For a scientist, a theory means an abstract model through which the phenomena available to experience can be related to each other in such a way that their nature and behaviour, as phenomena, seem to have been accounted for. But scientific theories only count if they have two kinds of clarity: the internal structure of the theory must be clear, and the reference to phenomena must be clear. These two conditions create the possibility of refutation, and refutability is the morality of science. If a theory does not predict what can be seen to be the case, or fails to predict what is the case, then we must eventually give up the theory and try another. We must also give up a theory if a simpler one explains the same phenomena. There is an aesthetics as well as a morality in science.

At first sight, architectural theories appear to be rather different. An architectural theory is usually presented as a set of precepts that, if followed, lead to architectural success. The prime aim of an architectural theory thus seems, not to be to explain architectural phenomena, but to guide design. We might therefore be tempted to conclude that architectural theories are normative rather than analytic – that is, they tell us how the world should be rather than how the world is – and are therefore not subject to the strict rules that govern scientific theories.

Although architectural theories do come in a normative mode, this by no means implies that they are not also analytic. On the contrary, it implies that they are. The only possible justification for a normative architectural theory is that the theory will work because this is the nature of architectural phenomena. Theories from Alberti to Le Corbusier, in fact, make profound and far-reaching assumptions about human nature, about perception, about behaviour, as well as about the nature of architectural order. It cannot be otherwise. All normative architectural theories are also, perhaps covertly, analytic theories.

The difference between scientific theories and architectural theories, then, is not a difference in type but a difference in clarity. It has never been possible to have architectural theories that have the two kinds of clarity – of internal structure and of reference to phenomena – that are the precondition for refutability. This is why architectural theories can be refuted by life, but not by analysis. One useful effect of space syntax is that it takes certain aspects of architectural theory a little way in the direction of the two kinds of clarity. The structure of reasoning is clear, and the reference to phenomena is clear. The propositions of space syntax can therefore be shown to be wrong – and theoretically wrong – by reference to evidence. Life is right, of course, and only life can eventually decide. But it is possible that, with theories that have the two kinds of clarity, more of life can be brought to bear on our theorising at the design stage.

Architecture as art – that is, as theoretical concretion

Does this mean, then, that the line between architecture as science and architecture as art needs to be redrawn closer to science? I do not believe so. We can call on the beautiful ideas of Ernst Cassirer on the relation between art and science:

> Language and science are the two main processes by which we ascertain and determine our concepts of the external world. We must classify our sense perceptions and bring them under general notions and general rules in order to give them an objective meaning. Such classification is the result of a persistent effort towards simplification. The work of art in like manner implies such an act of condensation and concentration . . . But in the two cases there is a difference of stress. Language and science are abbreviations of reality; art is an intensification of reality. Language and science depend on one and the same process of abstraction; art may be described as a continuous process of concretion . . . art does not admit of . . . conceptual simplification and deductive generalization. It does not inquire into the qualities or causes of things; it gives the intuition of the form of things . . . The artist is just as much the discoverer of the forms of nature as the scientist is the discoverer of facts or natural laws.[27]

Those of us who believe, as I do, that science is, on the whole, a good thing, accept that science is in one sense an impoverishment – although in others an enhancement – of our experience of the world, in that it cannot cope with the density of situational experience. It has to be so. It is not in the nature of science to seek to explain the richness of particular realities, since these are invariably so diverse as to be beyond the useful grasp of theoretical simplifications.

Science is about the dimensions of structure and order that underlie complexity. Here the abstract simplifications of science can be the most powerful source of greater insight. Every moment of our experience is dense, and, as such, unanalysable as a complete experience. But this does not mean that some of its constituent dimensions are not analysable, and that deeper insight may not be gained from such analysis.

This distinction is crucial to our understanding of architecture. Architectural realities are dense, and as wholes, unanalysable, but that does not mean that the role of spatial configuration, for example, in architectural realities cannot be analysed and even generalised. The idea that science is to be rejected because it does not give an account of the richness of experience is a persistent but elementary error. Science gives us quite a different experience of reality, one that is partial and analytic rather than whole and intuitive. As such it is in itself valuable. It needs to be accepted or rejected on its own terms, not in terms of its failure to be like life or like art.

It is, in any case, clear that the dependence of architecture on theories, covert or explicit, does not diminish its participation in Cassirer's definition of art. This is true both in the sense that architecture is, like art, a continuous process of concretion, and also in the sense that, like art, 'its aspects are innumerable'. But there are also differences. The thing 'whose aspects are innumerable' is not a representation but a reality, and a very special kind of reality, one through which our forms of social being are transformed and put at risk. The pervasive involvement of theory in architecture, and the fact that architecture's continuous concretion involves our social existence, define the peculiar status and nature of systematic intent of the architectural kind: architecture is theoretical concretion. Architects are enjoined both to create the new, since that is the nature of their task, and to clarify and improve the theories that tie their creation to our social existence. It is this dichotomy that makes architecture distinct and unique. It is as impossible to reduce architecture to theory as it is to eliminate theory from architecture.

Architecture is, thus, both art and science, not in that it has both technical and aesthetic aspects, but in that it requires both the processes of abstraction by which we know science and the processes of concretion by which we know art. The difficulty and the glory of architecture lie in the realisation of both: in the creation of a theoretical realm through building, and in the creation of an experienced reality 'whose aspects are innumerable'. This is the difficulty of architecture. And this is why we acclaim it.

Notes

1. B. Hillier, 'The nature of the artificial', *Geoforum* 16 (1985): 163–178.
2. N. Cross, *Developments in design methodology* (Oxford: Wiley-Blackwell, 1984).
3. H. Simon, *The sciences of the artificial* (Cambridge, MA: MIT Press, 1969).
4. H. Sanoff, *Methods of architectural programming* (Stroudsbourg, PA: Dowden, Hutchinson & Ross, 1977).
5. D. A. Schön, *The reflective practitioner: How professionals think in action* (New York: Basic Books, 1992).
6. R. Scruton, *The aesthetics of architecture* (London: Methuen, 1977), p. 234; B. Hillier, 'Quite unlike the pleasures of scratching: Theory and meaning of architectural form', *9H* 7 (1985): 66–72.
7. J. Hanson, *Encyclopaedia of Architecture* (New York: McGraw-Hill, forthcoming). Editors' note: There is no publication by this title with these publishers. The *Encyclopedia of American Architecture* by William Dudley Hunt and Robert T. Packard (New York: McGraw-Hill, 1995) was checked for this quotation, but it seems not to have been published.
8. H. Glassie, *Folk housing in Middle Virginia: A structural analysis of historic artifacts* (Knoxville: University of Tennessee Press, 1975), IV. The Architectural Competence, pp. 19–40, 21.
9. M. Mead, 'The transmission of culture through artefacts', in *Continuities in Cultural Evolution* (New Haven, CT: Yale University Press, 1966), pp. 83–106.

10. Scruton, *The aesthetics of architecture*, 43.
11. To which, incidentally, he has written the best introduction. See Roger Scruton, *A short history of modern philosophy: From Descartes to Wittgenstein* (London: ARK Paperbacks, 1981).
12. R. Descartes, *Principles of Philosophy*, Principle X, in *The Philosophical Writings of Descartes*, translated and edited by D. Murdoch, J. Cottingham and R. Stoothoff (Cambridge: Cambridge University Press, 1985), pp. 177–292, p. 259.
13. B. Hillier, and J. Hanson, *The social logic of space* (Cambridge: Cambridge University Press, 1984).
14. See Hillier and Hanson, *The social logic of space*, for a mathematical development of this concept.
15. B. Hillier, J. Hanson and H. Graham, 'Ideas are in things: An application of the space syntax method to discovering house genotypes', *Environment and Planning B: Planning and Design* 14 (1987): 363–385.
16. D. Irwin, 'The house the architect built' (MSc thesis, University College London, 1988).
17. For a discussion of intelligibility, see next section, below.
18. See B. Hillier, 'The nature of the artificial', *Geoforum* 16 (1985): 163–178; and B. Hillier, 'The architecture of the urban object', *Ekistics* 56 (1989): 5–21. It is suggested in these papers that there are three types of spatial law: type 1 are limiting laws that govern the constructability of usable spatial patterns, and give rise to the fundamental analytic properties, such as depth and rings; type 2 are laws through which social forms express themselves in space, for example through the integration of certain types of function and the segregation of others, as seen in the French houses; and type 3 are laws through which spatial forms produce specific social effects, for example the effect of axial patterns on movement and thus on patterns of the natural co-presence of people in space, as set out in the next section.
19. For example, Camillo Sitte and Gordon Cullen. Even the Krier brothers are strongly influenced by this tradition.
20. All the computer programs used here form part of the *Space syntax* software developed by N. Dalton of the Unit for Architectural Studies at the Bartlett School of Architecture and Planning, University College London, with funding from the Science and Engineering Research Council and our industrial partner, t^2 Solutions Ltd.
21. A full study of all open spaces in London was made for the Mansion House Square Project in 1984 and reported in 'A Proof of Evidence of the Public Inquiry'. A second study of open spaces in London, including some new ones, was made in connection with the Broadgate Development and made available in a report called 'The Broadgate Spaces'. (Both reports by Unit for Architectural Studies, Bartlett School of Architecture and Planning, University College London, 1989.)
22. An axial map is defined as the smallest set of straight lines that cover the open space of the area. See Hillier, 'The nature of the artificial'; and B. Hillier, J. Hanson, J. Peponis, J. Hudson and R. Burdett, 'Space syntax: A different urban perspective', *The Architects' Journal* 178(48) (1983): 47–63.
23. These issues are dealt with at greater length in a series of recent papers. For an examination of natural movement, see B. Hillier, A. Penn, J. Hanson, T. Grajewski and J. Xu, 'Natural movement: Or, configuration and attraction in urban pedestrian movement', *Environment and Planning B: Planning and Design* 20 (1993): 29–66. See also B. Hillier and A. Penn, 'Dense civilisations: Or, the shape of cities in the 21st century', *Applied Energy* 43 (1992): 41–66, for a discussion of the movement economy; and B. Hillier and P. O'Sullivan, 'Urban design and climate change', in *Proceedings of Conference C59*, edited by Solar Energy Society (London: 1989), on the urban design implications of both. (Editors' note: it has not been possible to definitively identify this last publication.)
24. In the real situation, the experimentation took a rather different form and was more oriented to pragmatic design ideas and constraints. The experiments here are therefore illustrative, not historical.
25. B. Hillier, 'Theory liberates: Intelligent representations and theoretical descriptions in architecture and urban design', inaugural lecture, presented at University College London, 13 June 1990 (London: University College London, 1991).
26. J. Hanson, 'Order and structure in urban space: A morphological history of the city of London' (PhD diss., University of London, 1989).
27. E. Cassirer, *An essay on man* (New Haven, CT: Yale University Press, 1944), p. 143.

16 Virtuous circles, building sciences and the science of buildings (1994)

The science of design improvement: Or, what kind of science of space can inform architectural and urban design? Introduction to 'Virtuous circles, building sciences and the science of buildings'

Kayvan Karimi

'Virtuous circles, building sciences and the science of buildings: Using computers to integrate product and process in the built environment' was published in *Design Studies* in 1994. This is 10 years after the publication of Hillier and Hanson's first book, *The social logic of space*, and two years before Hillier's monograph *Space is the machine* was published. The transition from a more theoretical approach in the former to an applicable methodology for design studies outlined in the latter is quite evident in this paper. Unlike his many other publications, it was written for a wider audience – spanning from academics to practitioners, planners and other decision-makers – that would need direct and compelling persuasion to adopt a design approach that is normally frowned upon by mainstream designers: a science-informed approach to design.

The primary underlying argument in this article is that design is essentially a sequential and iterative process. Perceiving design in this way dismantles the common misunderstanding of design as instant idea generation. By understanding design as a process, we can see that intuitive and objective patterns of decision-making are interwoven. This immediately leads to the further important argument that spatial design cannot be an instantaneous act by its very nature, but is in essence an evolutionary process, since it requires the integration of the different functional considerations that interact in the form of the building. Thus, spatial design lends itself to a process of iterative development that can incorporate 'user feedback' from 'user experience': a positive iteration loop, namely a *virtuous circle*. Hillier and Penn's initial reaction to their proposition is simple and rather common-sensical: 'Why not?' However, they soon provide a more compelling argument. Resolving the complexity of independent functions that interact seamlessly and have a knock-on effect on each other demands an approach which must be based on a 'science of building' that can provide the understanding of interrelated functions, to complement what we explore for each independent function through 'building sciences'. In other words, only a science of building can inform the virtuous circle of design.

So, how could a science of building, or indeed a broader science of the built environment, be developed and implemented? The article proposes that understanding the structure of space is the foundation for this science of building: 'space is what we use and sell' and 'it is the common ground of all our decisions'. This shifts the argument to how we model and analyse the spatial structure. Through the definition of person-geometry analytical units, an axial line, a convex space and an isovist, the article argues that the structure of space can be understood through the analysis of functional space in direct relationship with its users. This way of modelling embraces the geometry of space, but it is heavily based on what people do in the space.

The article explains how a space science can provide an advanced understanding of the built environment, through a series of city-wide and neighbourhood-wide analyses of the spatial networks in London. To spell out how the approach would work, three case studies are presented in detail: first, an analysis of a dysfunctional housing estate in London, aiming to understand why it does not perform well and how it needs to change; second, an analysis of a work environment to determine how a healthy office environment can be assessed and improved; third, the development of a theoretical model to link configurational and metric properties of the spatial network into one single model.

Looking back at these three case studies nearly three decades later, we can trace back many of the theoretical and methodological developments of the space syntax research field. The adopted methodology and rigorous analysis demonstrated for the sample case studies have also informed countless projects[1] in which analytical, evidence-informed consultancy has been provided for architectural, urban design and urban planning projects.[2] In that sense, the article seems to have been well ahead of its time – although some amount of foresight can be explained by the evident demand for such consultancy through the activities of the Space Syntax Laboratory at that time.[3]

The first case study, centred on the London spatial network, laid the foundation for the further development of fundamental space syntax theories such as 'centrality as a process' and the 'part-whole structure of cities',[4] as well as critical methodological developments, such as city-wide and regional, or even country-wide spatial network modelling.[5] It was also the beginning of the subsequent development of complex pedestrian movement/vehicular forecast modelling. Although 'Virtuous circles' does not directly address the design iteration loop, it has informed many other studies that used this approach to improve the design of housing estates and mixed-use developments in the UK and abroad.

The second case study in this paper has influenced many subsequent studies on complex buildings and the development of a science of buildings required for a virtuous circle of design. It has also inspired the development of further computer-based methodological approaches, such as the use of GIS to analyse space in tandem with multi-layered databases,[6] as well as analytical methods such as Visual Graph Analysis and vision-driven agent-based modelling.[7]

Despite being a theoretical investigation, the third case study in this article has been very influential in the further methodological development of space syntax. By proposing a method of bringing metric distance and integrating it with the topological modelling of the spatial systems, the case study paved the way for the development of complex modelling techniques, such as 'weighted urban models', 'weighted agent models' and, most recently, 'Integrated Urban Models (IUMs)'.[8]

Despite its relatively simple language and straightforward style in explaining phenomena, this article should be considered one of Hillier and Penn's most influential contributions to the space syntax field of research. More importantly, it serves as an essential gateway to how this type of research can be realistically and effectively used in the process of architectural and urban design. This is something that Hillier has always been passionate about and considered as one of the greatest virtues of space syntax: an enhanced 'virtuous circle' of design.

Virtuous circles, building sciences and the science of buildings

Using computers to integrate product and process in the built environment

Bill Hillier and Alan Penn

Originally published: Hillier, B. and Penn, A. 1994. 'Virtuous circles, building sciences and the science of buildings: Using computers to integrate product and process in the built environment'. *Design Studies*, 15, 332–365.

Introduction

Buildings and built environments must satisfy a range of functional criteria – structural, environmental, economic, social, organisational, visual and so on. These functional criteria are independent, in that they are nothing like each other, but are interactive, in that when you change a building to get one right, you may make something else go wrong. This creates two knowledge problems in the making and managing of buildings. One is integrating knowledge of the product and its functioning into a better understanding of the building as a complex whole. The other is integrating the process to create the virtuous circle of progressive product improvement through feedback from user experience. In this paper, we argue that these are aspects of the same problem and have a common solution. The solution starts from a very simple observation: that the different functional criteria affect each other only through the building. It follows that to see how they relate, we must therefore take a building-centred rather than a discipline-centred view of buildings and how they function. This means that we need techniques for modelling buildings which are sensitive to the different aspects of function. Such a technique is possible through spatial modelling, that is modelling the structure of usable space created by the building rather than the building itself. In this paper, we outline an approach to such modelling and its applications to the problem of product and process integration through building-centred studies. Once we do this, we develop a building-centred view of the forms and functioning of buildings which can then be used to integrate the process and most importantly to create the virtuous circle of feedback from user experience into progressive product improvement.

The problem: We have building sciences but no science of buildings

A constant feature of the built environment industry is that we seem to find it hard to develop 'virtuous circles' in which products – buildings and places – are progressively improved and made competitive by feedback from user experience. Of course, we have post-occupancy evaluations and recently the RIBA[9] has called for much more research based on how buildings are really used. But we cannot pretend that 'virtuous circles' yet exist, or even that they are in sight.

Why not? Is it that no one pays for it? Hardly. No one pays a manufacturer to do it. But manufacturers see product improvement, by feedback, as an economic necessity, because without it products are unlikely to remain competitive. So why does our industry not see this competitive necessity? We suggest that it is because our clients and users are as much in the dark as we are about what a genuine quality product is in our field. This is because our product is a very peculiar one, and we have not yet taken account of all its peculiarities in trying to establish 'virtuous circles'.

The problem is something like this. Buildings, as we all know, must satisfy a whole range of criteria: structural, environmental, economic, social, organisational, visual and so on. These criteria have two curious properties. First, they are so unlike each other that each must have its own technique of product evaluation and its own distinctive knowledge input into making buildings. As forms of knowledge, and methods, different functional criteria are independent. At the same time, when we change a building for one functional criterion, we often find it has knock-on effects on others, sometimes quite dramatic ones, so that we are in danger of getting something wrong as a direct result of getting something else right. Applied to buildings, the functional criteria are interactive.

This combination of independence and interaction is our problem. It is very difficult to define forms of knowledge or methods, either practical or theoretical, which are about the building as a whole. Because functional criteria only interact with each other through the building, we find it difficult to deal, with any rigour, with the interrelations between functional criteria. Since we develop building sciences and computer software broadly following the lines of the functional divisions, the problem is, if anything, made worse. We see the parts more clearly, the whole less clearly. As a result, building knowledge, as at present constituted, is an increasingly specialised set of subfields held together by a practice. We have building sciences, but we have, as yet, no science of buildings. We know enough to compete on time, on cost, and on some outward manifestations of quality. But we simply do not know enough about how buildings work as complex products to compete on the 'really good building' criterion and set up the virtuous circle.

So, what should we do? First, we should give up the dangerous illusion that interrelationships between these functional areas might be understood independently of the study of buildings, and seek instead to understand their interrelationships through the study of buildings. We propose that we now need to bring the building centre stage and give more attention to the development of a science of buildings to complement building sciences, one concerned with the actual and possible forms of buildings and how they interact with the various disparate functional criteria that we try to reflect in their form. We must do this because the only way that the functional criteria do interact is through the building.

In this development, the computer can and will hold centre stage because through the computer we can develop forms of modelling for buildings and built environments which are functionally sensitive and relatively intelligent. The basic argument is simple. If functional criteria, and the different forms of knowledge that they imply, do interact through the building, then it ought to be possible to model the building in such a way as to show how these interactions occur.

In this paper, we will outline one approach to intelligent building modelling in this sense and then show, through a series of live case studies, how we are using the building to work towards a more synthetic, multifunctional understanding of buildings and built environments. Computers are potential integrators, we will suggest, not just in the sense that we should interrelate software packages and standardise hardware, but in the sense that we can use the power of computers to synthesise the different forms of knowledge and understanding that built environments require, by using the building itself, and our ability to model it with more and more intelligence, as the synthesising framework.

Modelling built environments

How can we model a building or built environment in such a way as to make the model sensitive to function? The answer may at first seem surprising. We model the structure of space created by the building, inside and outside, as in Figures 16.1 and 16.2. This is not as bizarre as it may sound at first. Space is, after all, what we use in buildings. Walls make the space, and cost the money, but space is what we use and sell. Seen functionally, built environments are about the use of space created, by a building, inside and outside. Space is the common ground of all our decisions. Whatever criterion we are trying to achieve, our decisions eventually make a difference to how we use the form of the building to make a pattern of usable space.

But is there another more powerful and substantive reason for studying space. We have been able to show, through research (as we will see below), that the way

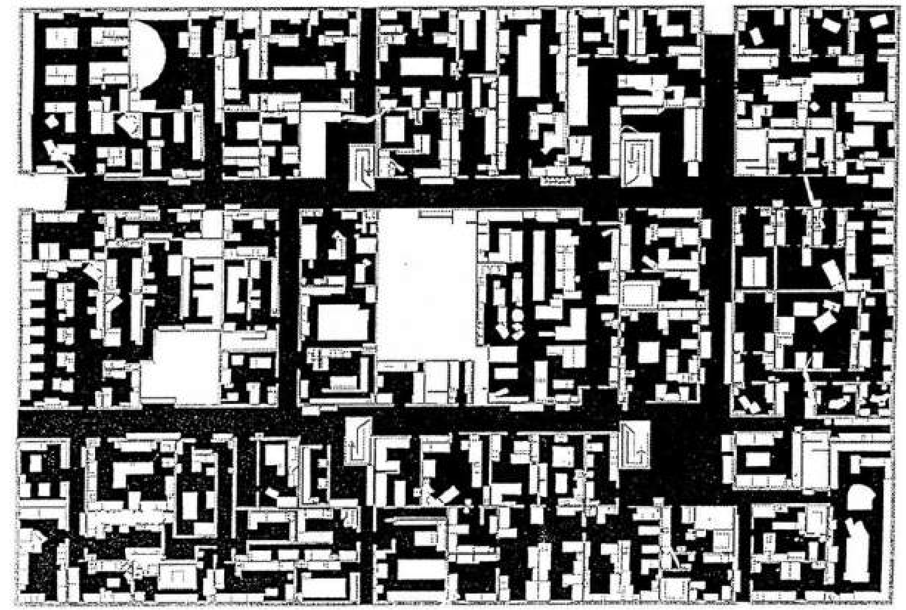


Figure 16.1 The structure of space in a building interior.

space is structured – that is, organised into a pattern – has, in itself, remarkable strong effects on how buildings and built environments function, so much so that we are convinced that this is at least one of the keys to how buildings and built environments work for organisations and communities.

How then do we model space? Evidently, we must do so in a way which is sensitive to the functional potentials of organised space. To grasp this, we must change our idea of space a little. We are accustomed to thinking of space as a kind of geometric shape, a room, a corridor, a street and so on. As soon as we think of space in this way then it becomes very difficult to imagine how there could be any relationship between a vacant shape and what people do, except the obvious constraints such as not being able to get enough people in the room. But the effects of space on organisations and communities are not at this level. They are at the level of the system of interrelated spaces. It is, as such, that we must learn to analyse spatial layouts.

To do this in a functionally intelligent way we must forget about simple geometry and first think spatially about what people do. When we do, we find that much human activity has what we might call a natural geometry. People move in lines, talk in convex spaces, see irregular shapes we call isovists (Figure 16.3). A spatial layout can be seen as offering different functional potentials. What is it like to move around in it? Does it have the potential to generate interaction? Can strangers understand it? And so on. All these questions are about the relationships of spaces, as functional potential, and how they are organised into a system.

Figure 16.2 The structure of open public space between buildings.

A layout can thus be represented as a different kind of spatial system according to which aspects of function we are interested in. For example, if we are interested in movement – and we should be because it is one of the most fundamental aspects of space use and much flows from it – then we can model a system of space as a system of lines, that is, as a matrix of all possible routes. For example, Figure 16.4 is a line map of all routes in London, approximately within the North and South Circular Roads. Figure 16.5a is a 'depth' analysis of the system of routes from Canary Wharf. Lines immediately adjacent to Canary Wharf are printed darkest, then, each time

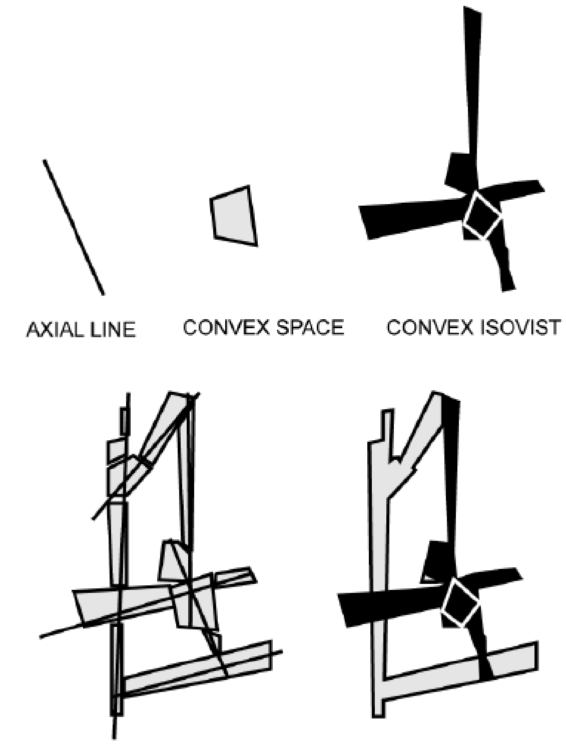


Figure 16.3 People move in lines, talk in convex spaces, see irregular spaces we call isovists. Superimposed, these elements create overlapping patterns.

you tum a corner moving outwards from Canary Wharf in any direction, you lighten the line a little. A very light line will then be very many lines 'deep' from Canary Wharf. The analysis pictures, if you like, the complexity of routes from Canary Wharf to the rest of London – not its metric accessibility but its ease of accessibility. Now, in Figure 16.5b, look in contrast at the same analysis from Oxford Street. It shows graphically how much greater the ease of access is to Oxford Street from the whole of London. Canary Wharf is deep from everywhere, Oxford Street is shallow from everywhere.

These pictures of London, from the point of view of each of its constituent lines, can be measured exactly through a quantity we call integration.[10] By this we can assign an 'integration value' to each line in the system, reflecting its mean depth from all other lines in the system, and comparing it to the theoretical maximum and minimum that it could have. We can then map these integration values from dark to light and produce an 'integration map' of the whole of a city, as in Figure 16.6. We can see that Oxford Street is London's strongest integrator because it is on average 'shallower' from all other lines than any other.

Figure 16.4 London, all routes line map.

Figure 16.5a London, point depth from Canary Wharf Tower. Darker lines have less depth.

Figure 16.5b London, point depth from Oxford Street. Darker lines have less depth.

Figure 16.6 London, integration map. Darker lines are more integrated.

Integration values in line maps are of great importance in understanding how urban systems function because it turns out that how much movement passes down each line is very strongly influenced by its 'integration value' calculated in this way, that is, by how the line is positioned with respect to the system as a whole.[11]

In fact, it is slightly more subtle and depends on the typical length of journeys. Pedestrian densities on lines can be predicted by calculating integration for the system of lines up to three lines away from each line (radius 3 integration), as shown in Figure 16.7, whereas car densities depend on higher radius integration because car journeys are, on the whole, longer and motorists therefore read the matrix of possible routes according to a larger-scale logic than pedestrians. We are currently, with SERC support, exploring the consequences of these results for how we design and manage the relationship between pedestrian and vehicular movement.

Now consider Figure 16.8, which is a radius 9 integration map of London, that is an integration map calculated to a radius equivalent to the mean depth of the most integrating line at radius *n* from the rest of the system. This is believed to be the truest map of the urban structure. It is actually quite remarkable, because although it has been made purely by a mathematical analysis of the matrix of line interconnections, not only can we predict movement with it, but also we can see how very powerfully the distribution of land uses seems already to have been affected by the structure of integration, and presumably, the effect they have had over time on the movement pattern.

These results suggest some quite deep relationships between different aspects of urban functioning and between the form and function of cities. Research now suggests that many, if not most, aspects of how cities work are influenced by this core relationship between the structure of the urban grid and the movement pattern it engenders: location of retail and housing areas, vulnerability to crime, social performance, even – more speculatively – economic performance.

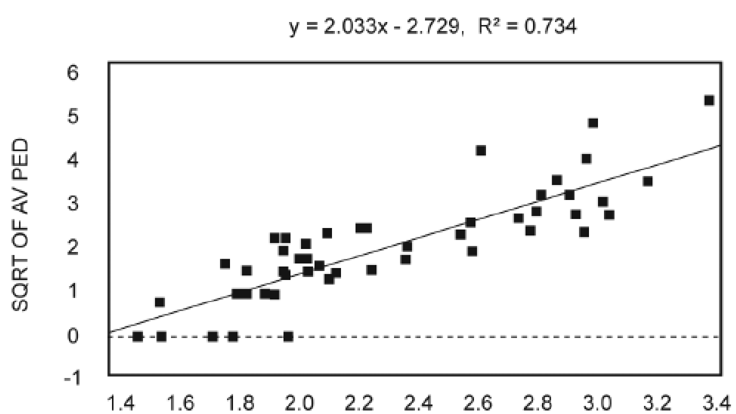


Figure 16.7 Correlation between integration radius 3 and rates of pedestrian movement.

Figure 16.8 London, radius 9 integration. Darker lines are more integrated.

Two important consequences follow from these results. First, we can use the combination of the analysed spatial representation of the built environment and the associated functional statistical correlates as an intelligent basis for integrating knowledge of different aspects of urban function. We can proceed visually by drawing dots on maps to see patterns, and statistically to draw out relationships between the spatial pattern and the different aspects of function, and between different functions.

Second, because the basis of our research investigations has been a spatial representation of the built environment itself, we have a functionally intelligent map which we can use immediately and directly in design mode. By drawing new proposals into our functionally intelligent map, and reanalysing the system with the new proposals, we can begin to explore the likely functional effects of spatial changes across a whole range of criteria from movement patterns to land use decisions, and from social performance to crime vulnerability. This 'space syntax' modelling is now being used on a large number of major urban design projects, including the recent, competition winning, Rogers' proposal for the new Central Business District for Shanghai.

But most importantly, by creating a means of integrating our understanding of function with our understanding of the built product, we create the basis for continuous monitoring of the relationship between the built environment and its functional performance, which does open up the prospect of creating the virtuous circle in which we learn from what we have done and feed it back into the next round of decision-making.

We will now examine this possibility in more detail through a series of live case studies. The three case studies are on different themes, illustrating something of the current range of application of these techniques, but a developmental theme runs through them. The first shows, at the simplest level, how we can use the spatial modelling linked to various types of functional outcome to both research into a problem, and also directly apply the results in design. In the second, we add a much larger range of functional data, while keeping the spatial model relatively simple. In the third, we develop a great deal more complexity in the spatial model, with the warning that, in this case, the work is in the early stages and not all these additional spatial properties are yet fully understood in terms of their functional implications.

Three case studies

Case study 1: Redesigning a bad housing estate

Our first case study is a common problem today: how to redesign a housing estate to make it an acceptable environment. The case study we will describe has a slightly curious history. We came to be involved through the BBC, who wanted to use it as a test case in a short programme they were making about our work for *Tomorrow's World*.[12] The North Peckham estate is an early 1970s estate built at the time when, in reaction against the tower and slab approach, our designers had sought to recreate traditional, urban quality by building 'streets in the air'. The North Peckham estate has probably one of the largest networks of upper-level walkways in the world. It is also one of the most notorious estates in the country, often featured as the place that no one dares to go to. Some years ago, the first fad for rehabilitation had led to lopping off storeys and adding pitched roofs – with no discernible benefits. This time, change is to be more radical. The local authority architects had produced a design to relocate the whole estate access to ground level, knocking down some buildings and adding others and, in general, reshaping the whole system of space to make it like a real place. The BBC asked us to meet the tenants and evaluate what was being done to their estate this time.

We did three things. First, we used the computer modelling technique to carry out a study of the estate, in context, and show exactly how it was going wrong: why it was dangerous, why space was being abused and why it didn't feel human. The design had literally generated a pathological pattern of space use by creating lacunas in the system of natural movement into which kids were moving unsupervised and forming gangs (Figure 16.9). Second, we then used the same computer model in simulation mode to explore the local authority proposal in the light of what we had discovered about how the estate was working and, in this way,

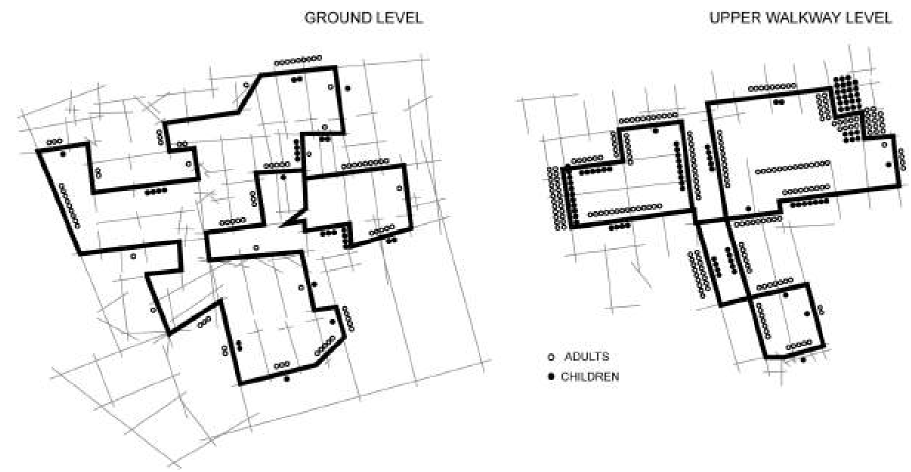


Figure 16.9 North Peckham estate, patterns of space use by adults and children.

Figure 16.10 North Peckham estate: (a) as built; (b) proposed development by local authority.

to show that the local authority scheme was half right but needed to be reworked to take into account the wider context (Figure 16.10). The result of this has been quite dramatic. The local authority are now looking at the whole area, and have related a whole series of estate proposals together (Figure 16.11). The existence of the modelling technique in which numbers are represented as colours had the effect of opening up a highly articulate public debate about what could, and should, be done in the light of what would and would not work. It democratised the design process, at least to some extent.

Figure 16.11 North Peckham estate, reworked local authority proposal.

The key thing about this case study is that by modelling space we are able to bring together all kinds of material on movement, surveillance, crime, social problems and vandalism into a common logic within which we can also discuss the design proposals. We are thus partly integrating knowledge of the product but also integrating the process by carrying knowledge forward from research and analysis into design simulations for use in refurbishment and management.

Case study 2: Designing healthy buildings

The second case study is a research project leading to design guidance, in general, rather than a single proposal. This study brings in a greater range of data, including environmental data, and greater integration of statistics with spatial analysis.

The background to this project is the questions that are currently being asked about the effect of buildings on health, not so much the 'killer buildings' with Legionella in the cooling towers, but the belief that buildings can produce chronic sickness over a period. Obviously, the environmental conditions that we create in the building through the HVAC system are at the heart of the debate. For decades we have designed for comfort. Is this the same as designing for health? If not, should we replace comfort criteria with health criteria, and if so, how should we do it?

To answer this question, we must understand much more than we do about the impact of buildings on health. The trouble is that the incidence of sickness, even in allegedly 'sick buildings', is not easily accounted for in a consistent way by specifically identifiable conditions.[13] A whole host of factors may, it seems, be involved over and above the basic questions of air quality and pollutants: spatial design and space use factors, management and organisational factors, work regime factors, communication factors, social and interpersonal factors, and so on – a multifunctional question if ever there was one. How do we disentangle all of these and isolate the effect of buildings on health?

To try to answer this, the Bartlett has set up a SERC/DTI[14] programme, in collaboration with the Welsh School[15] and various industrial and client bodies to carry out a case-study-based programme of research based on an integrative methodology that tries to take all of these factors into account. Figure 16.12 gives a partial picture of what we are trying to do. As the figure shows, we work with data in two modes: the visual and the statistical. Figure 16.12 is based on a detailed plan of a building on which we have superimposed several kinds of data. These include: details of occupant's health records and symptoms of ill health (part of an individual's responses are shown); long-term environmental measurements (illustrated by the line chart of temperature and CO_2 levels); observations of inhabitant space use patterns (shown by dots indicating the actual activities and locations of people); and analysis of the layout (illustrated by part of a line map showing the distribution of spatial integration).

All this is held in a framework formed by an analysis of the spatial layout of the building itself. Every location in the building is described numerically in terms of its local spatial characteristics and its relation to the building as a whole. People, either as individuals or as clusters (we increasingly use the term 'location groups'), are thus assigned values indexing their position in the building and the characteristics of that position. These numbers can then be compared by ordinary statistical means with numbers representing any aspects of function. In this way, the database keeps a precise record of the building in view and the precise relationship of people to the building every time an analysis is made.

Since the information is gathered at a precise and detailed level of individual locations within the building, it can be interrogated at many levels. The highest level of resolution is that of comparison between single individuals; the next level up is of groups of individuals who may be in spatially adjacent clusters (for example, a group of people sitting together and suffering from a diffuser malfunction) or a group of individuals who are in spatially diffused locations in the building, but suffer from the same problem (such as lack of fresh air, or over-isolation of workstations); the next level up from clusters is to compare across building floors or parts of floors; whilst the last level of resolution is to compare across whole buildings.

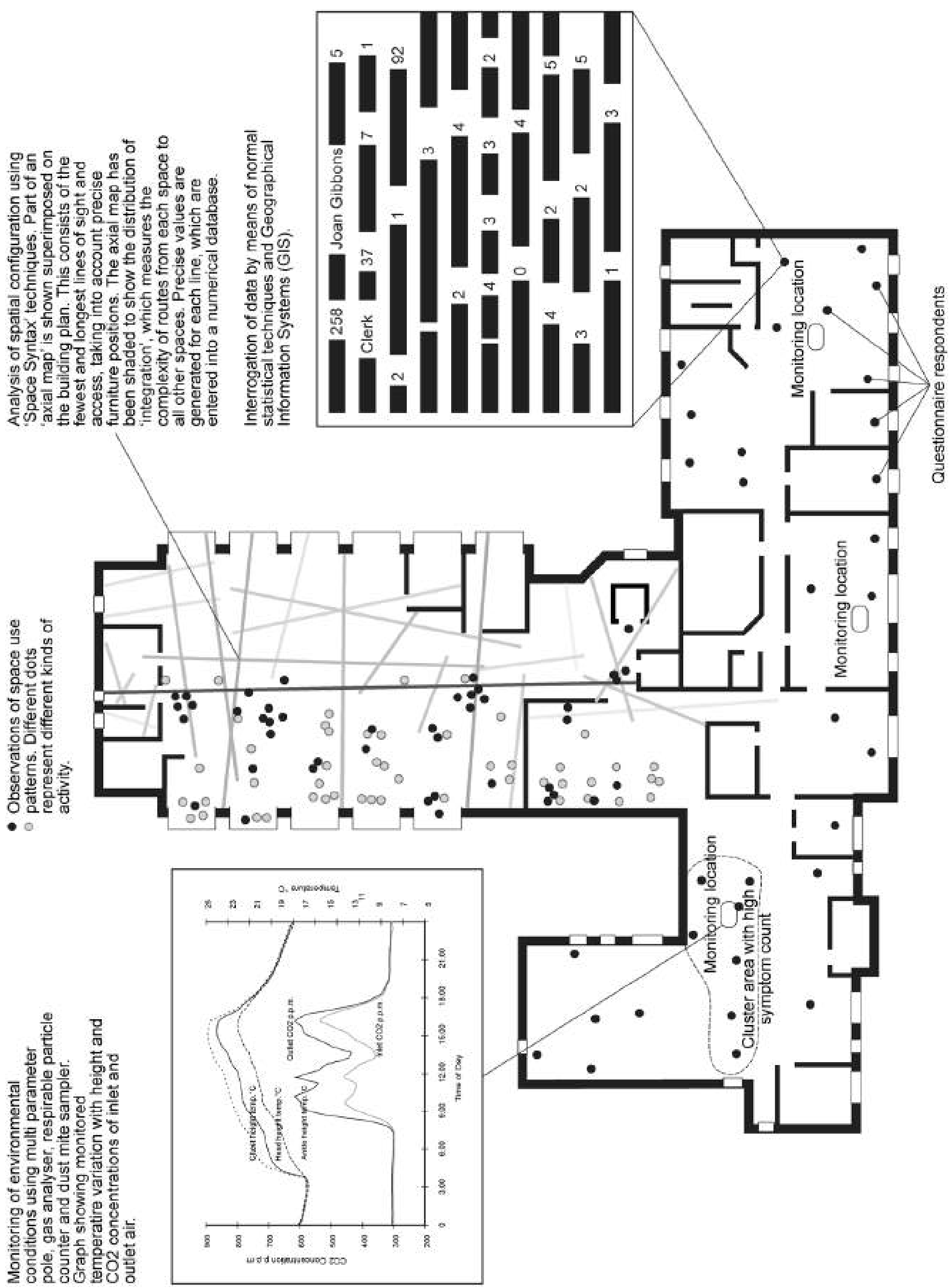


Figure 16.12 SERC/DTI LINK programme – new guidelines for the design of healthy office environments (the data in the diagram do not refer to actual research results of questionnaire respondents).

Using these methods, we have established that, in some buildings, there are concentrations of individuals who report high levels of ill health symptoms. These individuals report a high level of dissatisfaction with almost all aspects of the building environment, from air quality to lighting levels. However, actual measurements indicate that their conditions are no different from people reporting low symptoms. We believe that the underlying cause is a social and organisational process which is concentrating certain poor communicators in specific, isolated parts of the building complex, and these in turn are giving rise to the complaints. However, in other cases, we have found the opposite. In one instance the organisation has placed high-ranking individuals such as senior managers in over-exposed locations, where they are vulnerable to easy disruption and which, in turn, appears to be manifested in higher levels of stress and ill health.

Now what makes this analysis possible is the fact that some of the sets of numbers in the data table are numbers which describe the spatial characteristics of different locations and thus of the layout of the building from the point of view of that location. This holds the data not only visually, but also statistically, so that we never lose sight of how the building itself is involved in whatever processes were being identified through visual inspection and statistical analysis.

Although the project is only part of the way to completion, it may be possible to reach some general conclusions. We might say that the more general the symptom effect (indicated by indexes of complaint like the Personal Symptom Index), the more likely there is to be an organisational explanation. However, the more specific the complaint (such as instances of dry or watering eyes), the more likely it is to be a genuine effect of the environment in the building and be attributable to building services malfunction or poor design.

Case study 3: Intelligent analogues of cities

The third case study is different in kind again and has to do with one of the broadest and least tractable of issues facing the built environment industry: that of the economic, social and environmental 'sustainability' of cities. Even to effectively monitor and compare cities on sustainability criteria, whatever they might turn out to be, we must bring data on the physical and environmental performance of cities together with data on their economic and social performance, and relate both to some kind of description of the city. For example, energy consumption and pollution production depend, among other factors, on settlement patterns. Should settlements be dense or sparse, nucleated or dispersed, monocentric or polycentric, or a mix of all types? For research to give an answer, measurement data on environmental performance, and data on the implications of different behavioural assumptions (for example about the distribution of work and home) and 'knock-on' effects such as the economic, social and cultural consequences of spatial aggregation and disaggregation policies, must be related to descriptions of the physical and spatial form of cities which reflect the range of variation found in the real world.

We may begin by returning to Figure 16.8, which, we may recall, is a purely 'configurational' radius 9 model of London. Configurational models are already powerful predictors of movement but lack information on other key spatial attributes such as metric distance, area, density, plot ratios, shape, political boundaries and so on. In the past, we have brought such factors into data tables so that they could be statistically analysed in the same data frame as configurational and functional variables, as in the LINK project. We will now show how these spatial concepts can be expressed within the spatial model itself, thus creating a much more powerful basis for modelling cities.

The fundamental new idea is the integration of 'layered' representations of space into a single system for the purposes of analysis. For the purposes of illustration, we will use notional, simplified examples. Suppose we represent, as now, a street network as a series of lines or strips, as in Figure 16.13, and analyse their 'configurational' relations, for example to assign the pattern of 'integration' as previously so that strips (previously lines) are darker as they become more integrated, as in Figure 16.14. At this stage, no account is taken of metric distance, which, in some circumstances at least, seems likely to be an important variable.

However, we can add a metric dimension to the system by selecting an arbitrary module, say a 10-metre square, and linking modules into the pattern of grid as in Figure 16.15. We may analyse this on its own, as in Figure 16.16, but it is not of great interest, since it inevitably reflects the pattern of metric centrality in the grid. However, if we superimpose the line network onto the metric modular system and analyse the two layers as a single system, then the effect is to weight each line with a number of modules directly related to its length. The outcome of this 'length weighted' integration analysis is shown at both levels of the combined analysis: in terms of the modular units in Figure 16.17 and in terms of the 'line superstructure' of strip as in Figure 16.18. The strip level is much the same as previously, but the modular elements show an interesting, and very lifelike, localised structure in which greater integration is concentrated at the 'street intersections', with less integrated modules in the centres of links, away from the intersections. This immediately enables us to capture a functionally significant aspect of space organisation in a representation.

The relationship between metric area and configuration can be dealt with in an analogous way by underlaying convex elements with a two-dimensional modular layer, as in Figures 16.19a–f. In a–c we see how a simple system, in which four convex spaces of equal size and shape and the connections between them, are represented as a layer of modular elements with four convex elements and four strips for the connection superimposed. The two-layer system is then analysed. Whether we look at the result with the convex layer or the modular layer uppermost, the results will be a symmetrical distribution of integration dominated by the strips. In Figures 16.19d–f we give the convex elements different areas and underlay modular elements accordingly, so that each is now weighted by the number of modular elements it overlays. Analysis of the two-level system then

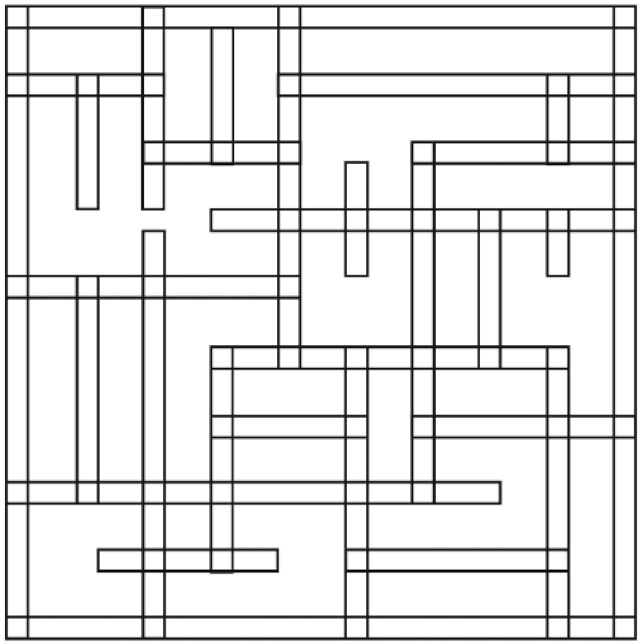


Figure 16.13 A street network as a set of intersecting strips.

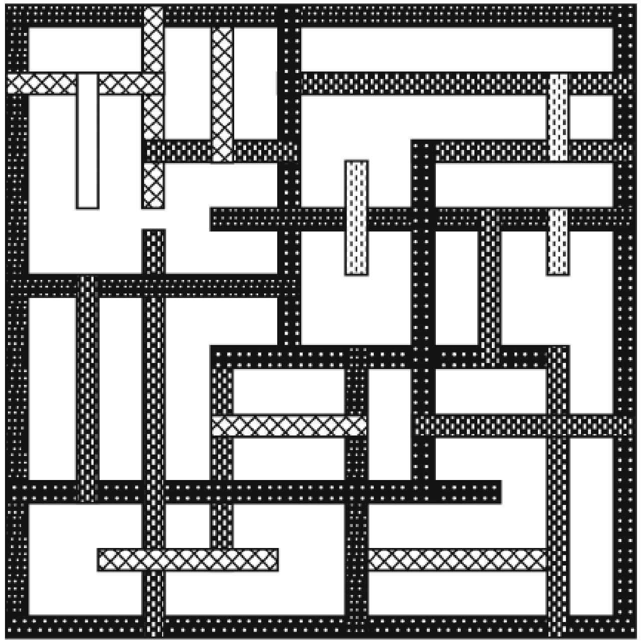


Figure 16.14 The pattern of integration of the network in Figure 16.13. Darker strips are more integrated.

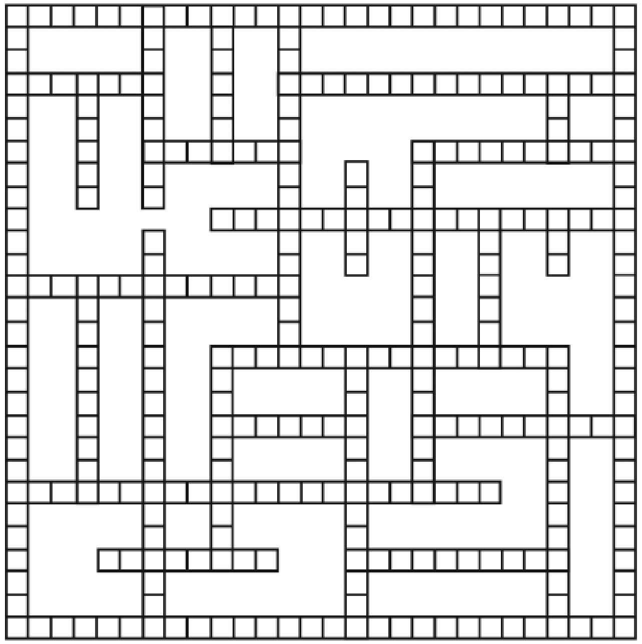


Figure 16.15 The street network of Figure 16.13 as a pattern of square unit tiles.

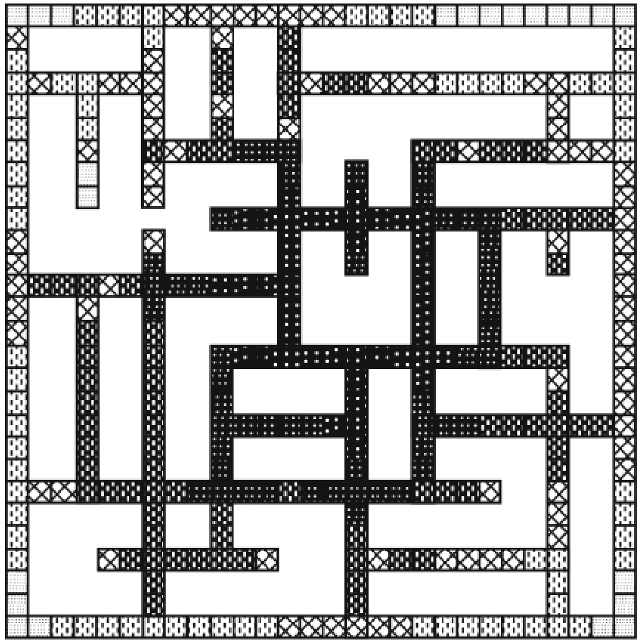


Figure 16.16 The pattern of metric centrality of Figure 16.15. Darker tiles are more directly accessible.

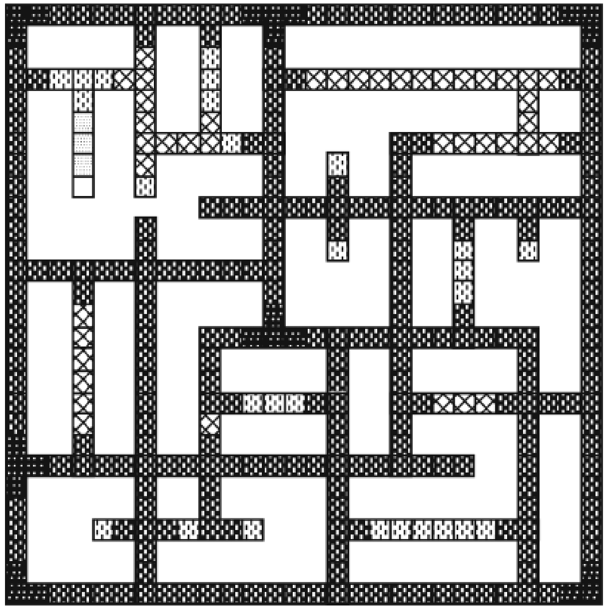


Figure 16.17 Length-weighted integration of the modular units of the network taking into account the linear elements represented in Figure 16.13. Darker areas are more integrated.

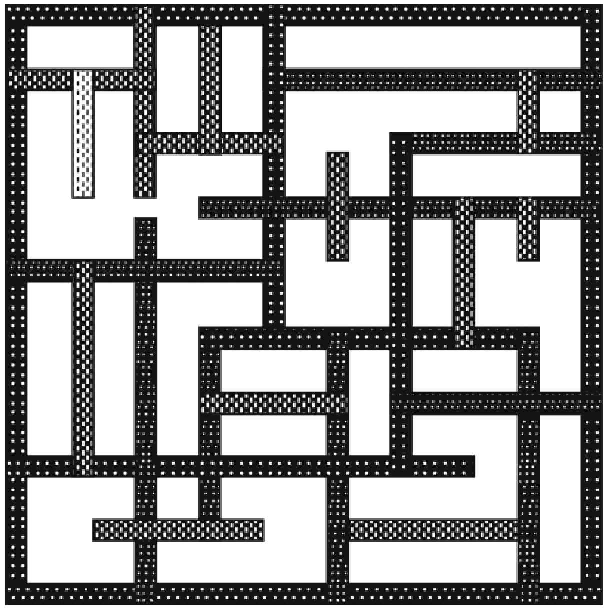


Figure 16.18 Length-weighted integration of the linear elements taking into account the modular units represented in Figure 16.15. Darker areas are more integrated.

shows that integration is drawn into the convex elements according to their area. Note, however, that the integration of the two smaller convex areas (on the top) is in the 'wrong' order. This is because the one on the left is closer to the largest-scale, convex area (bottom right) and this affects its own integration with respect to the rest of the system. Thus, the results show a combination of configurational effects and metric area effects. From this we can see that if we make a large and small square configurationally equivalent, in an urban system, then the large square will integrate more. Metric area, it turns out, is like distance, a property capable of expression as an aspect of configuration.

We may simulate the effect of plot ratios and densities by equally simple means. For example, if we wish to attach a building with a given number of floors to a street network, all we need to do is attach a convex space, the size of the ground area of the building, to the appropriate position in the street system, then overlay on that a convex element of each floor, making sure that each element above the ground is detached from the street and only connected through the ground layer as it would be in real life. This will not appear visually as a three-dimensional structure, but it will exactly represent the addition of above ground floor space to the urban system.

We may now build a model of an urban system in the following way. First, we divide the city up into an arbitrary number of areas and represent them as non-contiguous polygons. These may be as small or as large as we need, according to the level of resolution required by the research question. The polygons may be based on political boundaries, like wards, administrative boundaries like enumeration districts, segments defined by an arbitrarily fine grid, or they may be defined by objective morphological properties of the built environment. These polygons representing areas are the fundamental units of analysis for the technique.

Figure 16.20 shows an imaginary, simplified case. The street network of the city (or part-city) is then superimposed on the patchwork of polygons so that each polygon is linked into the urban system by all the streets or part-streets that pass through it or alongside it, as in Figure 16.21. This two-level, spatial system is then analysed 'configurationally' to find the pattern of integration in the whole system. The outcome is shown in Figure 16.22. Evidently, the street pattern will tend to dominate the area polygons simply because the streets are connectors. However, the street system can then be 'peeled off' the polygons, as in Figure 16.23, leaving a pattern of polygons with their spatial characteristics in relation to the city area around them, and to the city system as a whole, recorded as a set of numbers.

This basic process of linking areas together, by the street network, in a single configurational model, is the basis of what we call an 'intelligent urban analogue' model. Once this is established, we can then complicate the model in all the ways we have previously described. For example, we can underlay the street network with metric modules so that the analysis of the street system takes distances into account. We can underlay the polygons with metric modules so that the metric

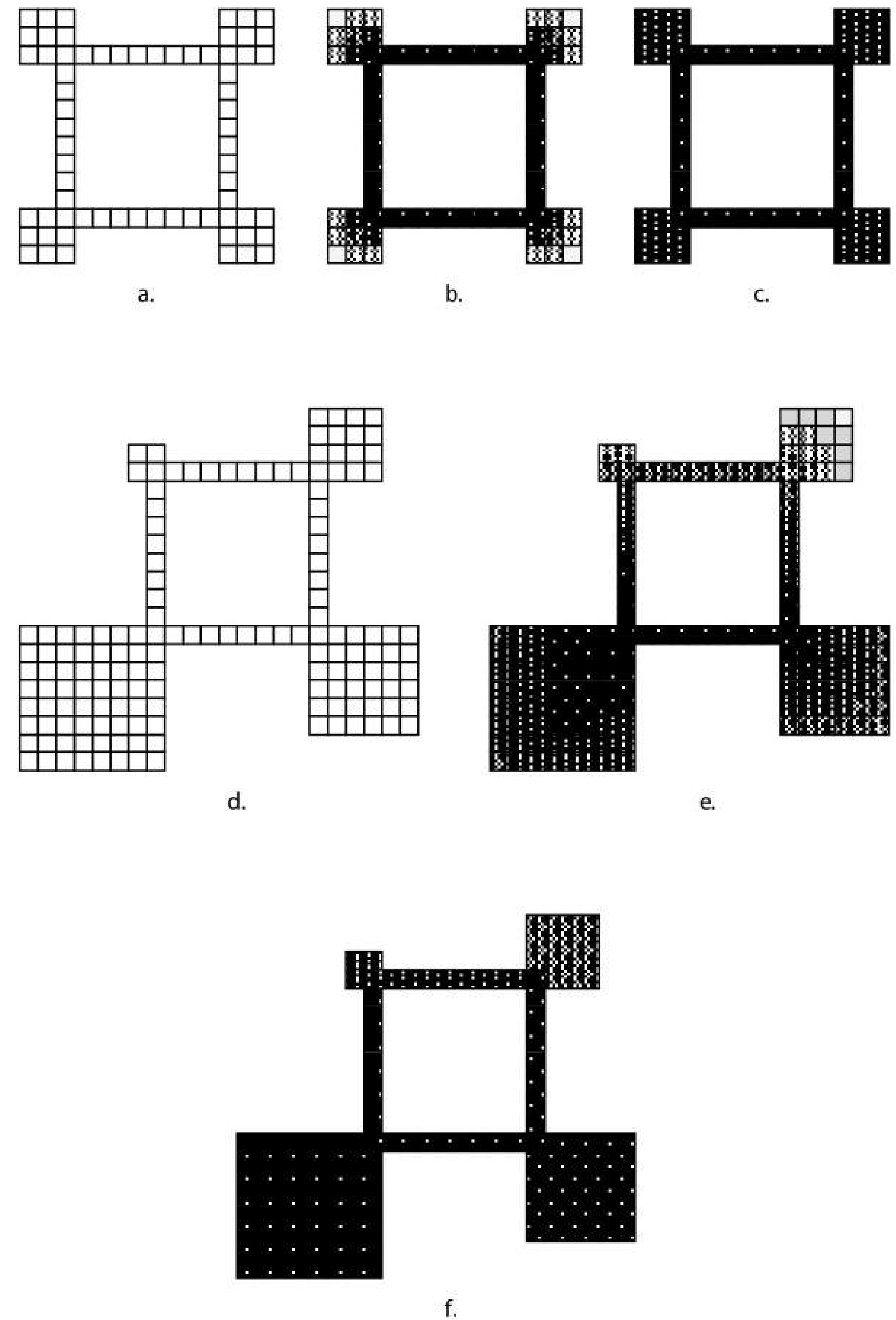


Figure 16.19 Two shapes (a and d) analysed for area-weighted integration, represented according to modular tiles (b and e) and according to linear strips and larger convex elements (c and f). Darker areas are more integrated.

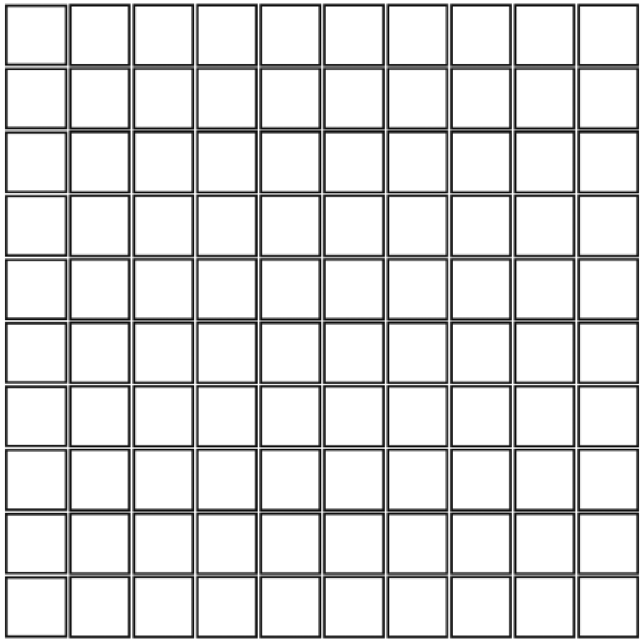


Figure 16.20 An imaginary city or part of a city represented as a number of areas.

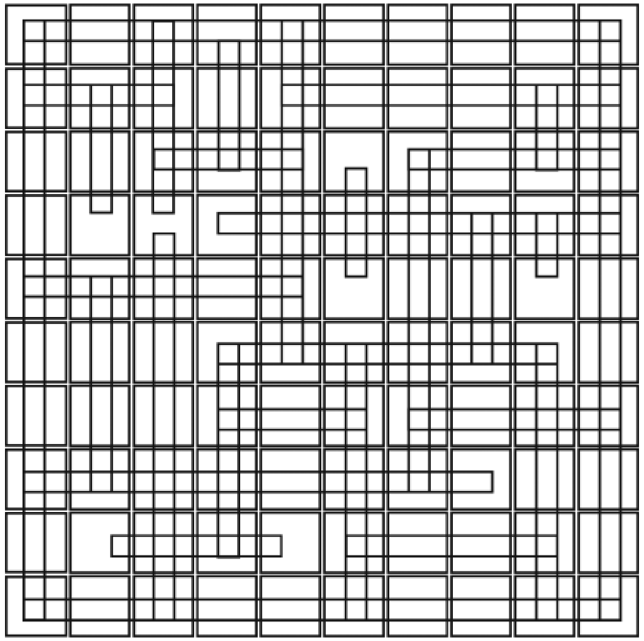


Figure 16.21 The areas represented in Figure 16.20 with superimposed strips representing streets.

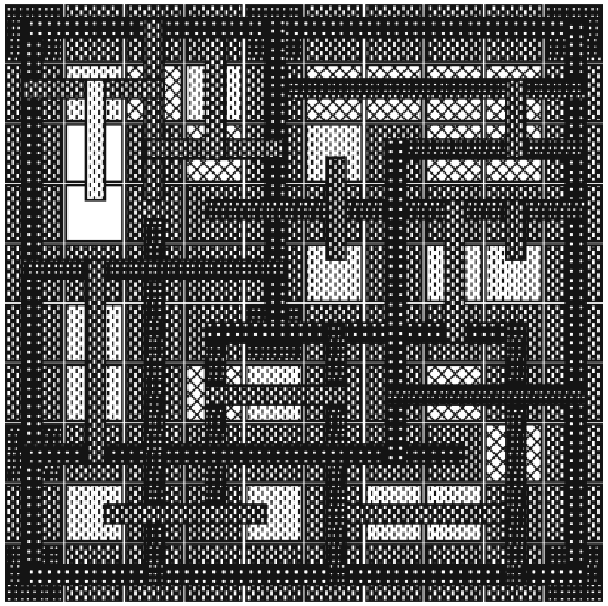


Figure 16.22 Area-weighted integration taking into account the areas and streets shown in Figure 16.21. Darker areas are more integrated.

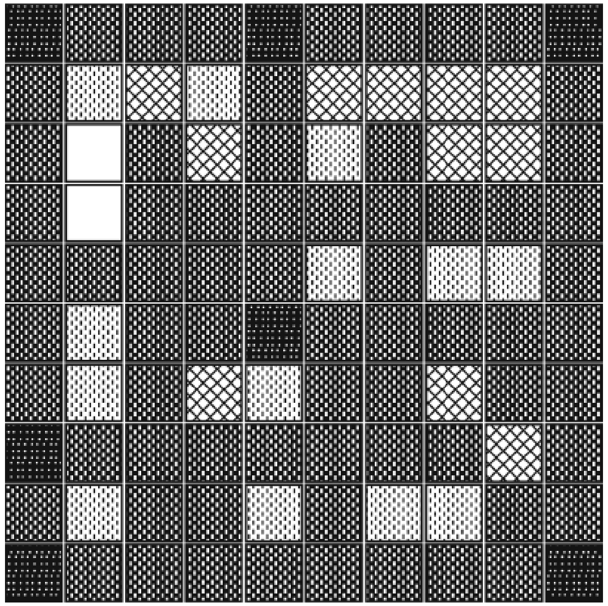


Figure 16.23 The distribution of integration of city areas after removing the strips representing the streets.

area of a polygon is taken into account. We can superimpose layers on the polygons representative of the ground floor space.

There is also an easy way of further disaggregating any model from the level of resolution originally selected. Each of the original area polygons can be, itself, subdivided into much smaller polygons and analysed as before. This more localised analysis will give a much richer and denser picture of the detailed characteristics of the area. These may then be fed into a larger-scale model as more detailed environmental descriptors. There is no reason, in fact, why both levels of the model should not be analysed as a single system. The principal barrier would be computing time. In our experience, adding a new level of fine structure to an existing model leaves the larger-scale picture more or less intact provided that the disaggregation is done uniformly and not confined to particular regions (which would lead to bias in the model).

At the other end of the scale, we may also derive new measures of the most macro properties of the city system, such as shape, and shape loaded with different densities in different regions. This can be done by simply linking the area polygons together and analysing the distribution of integration in the system without the superimposed street system. Shape will be indexed by the degree and distribution of integration and can be shown both by direct graphical representation of the city system or by statistical representations such as frequency distributions, or simply by numbers. The effects of weighting shapes by loading different regions with higher densities can be explored by simply overlaying the spaces representing the additional densities into the relevant polygons of the contiguous polygon system, then proceeding, as before, as in Figure 16.24, which shows the distribution of integration – and therefore the distribution of mean distances from point to all others – in a hypothetical urban system with two centres of distance apart. By varying the pattern and density of centres we can explore their effects on total distance travelled, all other things being equal, in a different kind of three-dimensional urban system. The effects of other, nearby settlements can also be investigated by simply adding them as extensions to the model.

The numerical data resulting from the analysis of the urban system can then be used in a number of ways. Firstly, and most obviously, the parametric descriptors for the polygons, resulting from analysis and reflecting as they do the position and configuration of each 'finite element' in the city system as a whole, then become the frame for other kinds of data which can be assigned as descriptors to the polygons. This can be done with any functional variable that can be numerically indexed for that area, such as population densities, pollution levels, traffic movement, pedestrian movement, unemployment rates, crime rates, council tax banding and so on. Because spatial and other descriptors are now all in numerical form, simple statistical analysis can begin to reveal patterns.

Second, and in our experience no less usefully, the distribution of any property may be represented graphically in the urban system as a visual distribution of that property in the city system. This means, in practice, all the visualising and

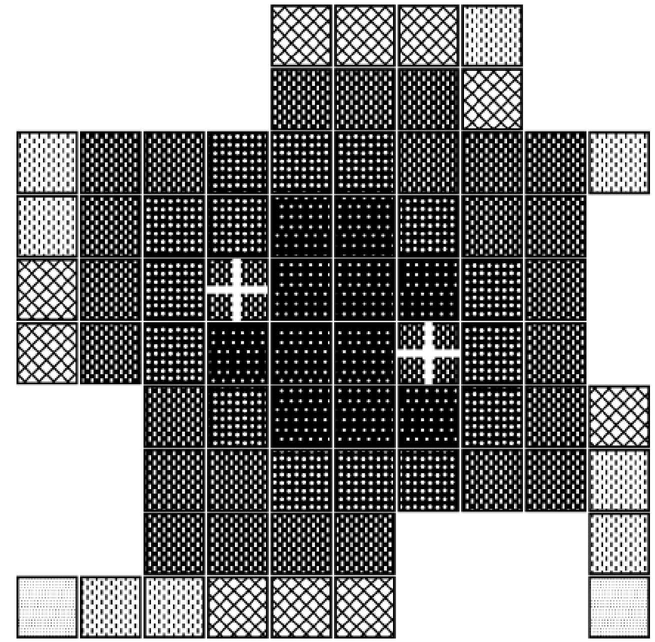


Figure 16.24 An urban system comprising a patchwork of areas is analysed as before, but after taking into account the density of development in two centres marked by crosses.

cartographical potentials that have been developed in the past few years through 'geographic information systems' can be interfaced with, and potentially bought within the scope of, an analytic model with proven ability to link morphological and functional properties of built environment systems in a predictive way.

And third, perhaps more contentiously, we think it may be possible to represent nonspatial attributes as a part of the configuration itself. For instance, job or population densities derived from census data and attributable to enumeration district polygons could be represented as numbers of modules attached to each polygon. The distribution of integration of the underlaying street network would then reflect differential distributions of the main trip origins and destinations in the city. Of course, this type of analogue model is only just under development. The incommensurability of spatial and nonspatial attributes may prove to be insurmountable. The proof will only come when we test its efficacy in accounting for empirical characteristics of city function. However, we believe that a system of this sort holds much potential for integrating the representation of a wide range of spatial and nonspatial attributes of real urban systems within a single coherent framework, and it is exactly this which is needed if issues as all-embracing as urban sustainability are to be tackled.

Conclusions

There is a common theme to all these studies. By using the computer to integrate knowledge of the product and its functioning/performance on the basis of an intelligent description of the building or built environment, we are able to use a common representation and understanding to integrate stages of the process. In the first case, we integrate from community consultation, through local authority planning and housing management to design and build. At the same time, we provide a tool for long-term monitoring so that, as time goes on, the effects of successive changes can be continuously reviewed. A virtuous circle can, in this way, be established in local authority housing design and management. None of us need reminding how much the absence of such a framework cost us during the heyday of local authority house building.

In the second study, we again integrate knowledge of the product and its functioning to link between occupancy and organisational studies, health studies, environmental monitoring and systems evaluation through to design guidance. Again we leave a tool for the facilities manager to monitor the progress of the organisation in terms of how it is getting on with its building. In the third case, we are explicitly bringing together different kinds of knowledge of urban performance on the basis of a description of the city in order to begin to construct the database we need on the complex and interrelated issues that go under the heading of 'sustainability'.

The essential message of all these studies, and of this paper, is that by integrating the product – and probably only by integrating the product – we integrate the process. It is only through a synthetic building-centred understanding of the objects we create and operate that we can create the virtuous circle of product improvement through feedback from user experience, whether we obtain this from continuous monitoring or from one-off studies.

In this, the computer must play a central role since it forms the means to integration of radically different kinds of functional relationship through a single model, as well as the means of providing quite different views of the models to which the different agencies, in a highly fragmentary industry, are accustomed. At the centre of this is the common representation of the one thing that unifies the whole field – the built and spatial object. This is the thing that designers manipulate and that organisations and people use. This, we suggest, is where we should now concentrate our attention.

Virtuous circles

An afterword

Alan Penn

Revisiting a paper after nearly 30 years can be enlightening. There are ideas in it that you had forgotten but that seem remarkably prescient, and others that on re-reading seem decidedly opaque. In this note I will try to give some context to help readers to understand where the 'Virtuous circles' paper came from and what, with the benefit of hindsight, I see as its legacy and interest.

This paper was written as a part of the final reporting on a Science and Engineering Research Council grant, SERC GR/H18647, with the snappy title of 'Graphical knowledge interfaces: Configurational models and precedent databases for built environment design teams'. This was a project which ran for two years between 1991 and 1993 at a value of £141,313. It was Bill Hillier's final project as a Principal Investigator with co-investigators Russell Winder from the Department of Computer Science, David Chapman from the Department of Photogrammetry and Surveying, both at UCL, and I, who had just made a move as a recently appointed lecturer from working under Peter Hall in the Bartlett School of Planning to working under Bill in the newly formed Bartlett School of Graduate Studies.

The grant was in response to a call for proposals developed by Professor James Powell of the University of Salford for the SERC Built Environment subject panel aimed at 'informing technologies for construction, civil engineering and transport'. Many of the key ideas for the proposal were developed by Nick Sheep Dalton, then the newest member of the Unit for Architectural Studies' research team and its computer guru. As the researcher being paid on the grant in those days, he could not also be a co-investigator, but the proposal was largely his idea and the team, including both Russell Winder, a database expert, and Dave Chapman, a GIS specialist, were brought in by Sheep.

There were several elements to the proposal. The first was based on a critique of the current state of GIS, which at that time was essentially a cartographic technology, lacking network analysis and statistical functionality. We would develop the analysis of configurations as relational systems consisting of networks of nodes. This would allow the calculation of graph measures to describe the shape and connectivity of the configuration. Second was the use of statistical representations such as scattergrams to examine both the properties of single spatial systems and whole families of these. This was the 'graphical knowledge interface' of the title. Third was the construction of 'precedent databases' of cities and urban systems. Bill and I had already been studying the ways that different configurational variables varied across large samples of cities and had used this, for example, to verify that the D-value relativisation was effective in removing the effect of system size, and to investigate the way that the properties of intelligibility and synergy varied across different urban forms. This work was extremely tedious and error prone, and we had a hunch that there was much more to be learned from a large and consistent dataset of urban form.

The 'graphical knowledge interfaces' program was to be developed on a NeXT Cube (Steve Jobs' 'next' enterprise after being booted out of Apple in 1985). In fact, for Sheep and me the main reason for writing the grant application was to get a few NeXT Cubes to play with. Quite apart from being extremely black and stylish (remember all tech in those days only came in prosthetic beige), we had heard that Tim Berners-Lee was using one to write something called a 'web browser', and it seemed like a cool thing to have. The NeXT Operating System (NeXTSTEP) was based on Unix and C++ and came with a very elegant graphical user interface builder. Sheep had already developed a graphically based analysis program on the Apple Mac for the 1989 'Vision of Britain Exhibition' at London's Victoria and Albert Museum and had ideas about how these kinds of interactive graphical programmes could be turned to knowledge generation. This was his forte, supplying Bill with 'tools for thinking with' and for 'the creation of phenomena' (to pinch Ian Hacking's epithet[16]).

This paper is a case in point. The all-singing and dancing NeXT program had expanded and expanded in functionality to the point where it was getting quite hard to handle. Sheep went away one weekend with his Mac and a new application called SuperCard and came back with a programme called 'Pesh'. Sheep recalls:

> 'I noticed it [SuperCard] had a nice feature in the scripting language where you could see if anything intersected anything else. So, a line intersecting a circle, a polygon intersecting a square etc. It made it easy to write Pesh in a day or two (peehs was sheep backwards). It was like writing a love poem in crayon, amazing but it worked. Bill saw the potential as a Thinker Toy.'

What Bill saw was that this tool let him play and experiment with relational properties of multiple networked objects. You could just draw them and copy and paste, and press calculate, and it would colour up all the graphical objects according to mean depth of each to all others. This is a global measure of the graph and something that is hard to be intuitive about until one has an interactive tool with visual feedback that allows you to train your intuition. And this is what Bill did.

Pesh became Bill's calculator for exploring configuration and presaged several years of his most creative thinking, giving rise to many of the key ideas that ended up in *Space is the machine*.[17] This paper may appear somewhat opaque, but it marked a key stage in Bill's way of thinking and the use of computing in generating theory and underpinned the fundamental concept of universal distance as well as partitioning theory.

Alan Penn, UCL, December, 2023

Notes

1. Space Syntax Limited (www.spacesyntax.com), a UCL spin-out company, is a consultancy firm founded by Bill Hillier and colleagues in 1989. The firm has worked on thousands of UK and international projects, spanning from small buildings to city-wide and regional planning. Bill Hillier participated actively in consultancy projects until his very last days.
2. P. Buchanan, 'Space syntax and urban design', in *Norman Foster: Works 3*, edited by D. Jenkins (Munich: Prestel, 2007), pp. 178–187; B. Hillier and T. Stonor, 'Space Syntax: Strategic urban design', *City Planning Review* 59 (2010): 7–11.
3. From the early 1980s onwards, the Unit for Architectural Studies at the Space Syntax Laboratory was invited to participate in myriad projects involving local government and scientists alike. See examples of this in B. Hillier, J. Hanson, J. Peponis, J. Hudson and R. Burdett, 'Space syntax: A different urban perspective', *The Architects' Journal* 178(48) (1983): 47–54, 59–63, and B. Hillier and A. Penn, 'Visible colleges: Structure and randomness in the place of discovery', *Science in Context* 4 (1991): 23–49.
4. B. Hillier, 'Centrality as a process: Accounting for attraction inequalities in deformed grids', *Urban Design International* 4 (1999): 107–127.
5. See for example the OpenMapping project for the UK: https://spacesyntax.com/project/openmapping-gb.
6. J. Gil, T. Varoudis, K. Karimi and A. Penn, 'The space syntax toolkit: Integrating depthmapX and exploratory spatial analysis workflows in QGIS', in *Proceedings of the 10th International Space Syntax Symposium*, London (Space Syntax Laboratory, The Bartlett School of Architecture, UCL, 2015), 108.1–108.18.
7. A. Turner, M. Doxa, D. O'Sullivan and A. Penn, 'From isovists to visibility graphs: A methodology for the analysis of architectural space', *Environment and Planning B: Planning and Design* 28 (2001): 103–121; and A. Turner and A. Penn, 'Encoding natural movement as an agent-based system: An investigation into human pedestrian behaviour in the built environment', *Environment and Planning B: Planning and Design* 29 (2002): 473–490, respectively.
8. P. Ferguson, E. Friedrich and K. Karimi, 'Origin-destination weighting in agent modelling for pedestrian movement forecasting', in *Proceedings, Eighth International Space Syntax Symposium*, Santiago, Chile, 2012; and K. Karimi, 'Space syntax: Consolidation and transformation of an urban research field', *Journal of Urban Design* 23 (2018): 1–4, respectively.
9. RIBA, *A Research Policy for the Profession* (London: Royal Institute of British Architects, 1991).
10. B. Hillier and J. Hanson, *The social logic of space* (Cambridge: Cambridge University Press, 1984).
11. B. Hillier, A. Penn, J. Hanson, T. Grajewski and J. Xu, 'Natural movement: Or configuration and attraction in urban pedestrian movement', *Environment and Planning B: Planning and Design* 20 (1993): 29–66.
12. Editors' note: This episode of *Tomorrow's World* was broadcast 5 June 1993 on BBC 1 London, directed by Dana Purvis.
13. G. J. Raw, *Sick building syndrome: A review of the evidence on causes and solutions*, HSE Contract Research Report No. 42/1992 (London: HMSO, 1992).
14. 1993–96 Intelligent Architecture research project funded by: DTI/SERC, GR/J 33609 Advanced Technology Programme £638,938 (total value £1.48m). PI: Alan Penn, Co-Is: Prof. Treleavan, UCL Dept. of Computer Science, & Prof Bill Hillier (eds.)
15. The authors are referring to the Welsh School of Architecture at Cardiff University (eds.)
16. I. Hacking, 'The creation of phenomena', in. *Representing and intervening: Introductory topics in the philosophy of natural science* (Cambridge: Cambridge University Press, 1983), pp. 210–219.
17. B. Hillier, *Space is the machine: A configurational theory of architecture* (Cambridge: CUP, 1996). Incidentally, the title of the book was also due to Sheep.

17 A theory of the city as object (2002)

How spatial laws mediate the social construction of urban space. Introduction to 'A theory of the city as object'

Meta Berghauser Pont

In 'A theory of the city as object', Hillier focuses on the processes that generate settlement form, namely the configuration of urban grids, and the laws or mechanisms behind these processes. The article thus extends earlier discussions on how the urban grid, through its effect on movement, shapes and is shaped by land use patterns and densities.[1]

The article is a typical example of Hillier's approach, starting from a proposition that is sustained through logical arguments, supported by case studies, experiments and mathematical analysis. Exploring urban morphology using more abstract and mathematical methods is relatively new and has been developed mainly during the past 30 years, as Batty points out in the foreword of *The mathematics of urban morphology*.[2] He highlights the work of Martin and March at the Centre for Land Use and Built Form Studies from the 1970s as pioneering the quantitative approach, mainly focusing on the geometric principles of form. Hillier's work, including this article, fits in this approach of describing the spatial laws behind the processes of growth or transformation that offer architects and urban designers the insight and means to, to quote Moudon, 'practice their art with its due precision'.[3]

The article opens with the observation that cities and their street configuration show an 'invariant pattern of local differences and global similarities'. Variations are found in the geometric properties of the urban grid such as line lengths, angles of incidence and intersection characteristics. For instance, Atlanta (USA) has a number of very long lines approximating the radius of the system and a strong right-angle discipline, while the longest lines in Hamedan (Iran) are only a fraction of the radius of the system and with a whole range of angles of incidence. Despite these geometric differences, striking similarities are observed in, for instance, the distribution of line lengths and block sizes. Thus, despite the geometric differences, both cities have a certain similarity of structure, with a hub and spokes in all main directions (Hillier's 'deformed wheel'),[4] smaller and more convex blocks near the centre and larger along the spokes.

It is then proposed that socio-cultural processes mainly affect the local geometric properties and thus the differences between settlements, whereas micro-economic processes lie behind the similarities found in these different settlements, resulting in the emergent global structure called the deformed wheel. Both, Hillier argues, are driven by the same spatial laws to generate or restrain potential movement in the system. The micro-economic process operates to maximise integration and thus co-presence in public space, while the socio-cultural process constructs the less integrated background of mainly residential space.

The spatial laws (or we could call them, in line with the terminology of Martin and March, geometric principles) are essentially clarifications of the above and demonstrate how spatial interventions affect distance and thus movement and co-presence, using logical arguments supported by simple geometric models and mathematical methods. Hillier introduces two laws here to which I suggest adding a third, following the propositions in Martin's *The grid as generator*.[5] My point is not to criticise but to highlight the unexplored potential of combining the lines of thought of Hillier, Martin and March.

The first geometric principle outlined by Hillier, the 'Law of Centrality', proposes that to minimise the increase in distance in the whole system as the city grows, a new object should be placed close to one object and far from the other. This results in settlements with long and short lines rather than lines of similar length. The law thus generates spatial inequalities with an integrated global structure of public space and a less integrated background of mainly residential space. The second geometric principle, the 'Law of Compactness', proposes that compact urban forms or blocks generate more local interaccessibility than elongated blocks of equal area and serve to support a sociocultural process that results in variations in local geometric properties. As explained above, this implies a third principle, a 'Law of Density',[6] whereby centrally placed forms or buildings generate taller buildings to achieve the same area of floor space than would be the case with peripheral placement. One could argue that this law is of a different kind, but it generates a denser core and a sparser periphery in the same manner as the law of centrality generates spatial inequalities in the global structure.

Hillier's article ends with a reflection on the reason why the topological analysis of urban grids is better in its capacity to explain the movement patterns despite the fact that the deformed grid is the outcome of geometric processes. The role of visibility in how we cognise and interpret our surroundings is suggested as an explanation and 'laws of behaviour' could thus be the reason for the success of configurational models because 'they reflect the world we see rather than the world of distance and mass'.

A theory of the city as object

Or, how spatial laws mediate the social construction of urban space

Bill Hillier

Originally published: Hillier, B. 2002. 'A theory of the city as object: Or, how spatial laws mediate the social construction of urban space'. *URBAN DESIGN International*, 3–4, 153–179.

Introduction

A series of recent papers have outlined a generic process by which spatial configurations, through their effect on movement, first shape, and then are shaped by, land-use patterns and densities.[7] The aim of this paper is to make the spatial dimension of this process more precise. The paper begins by examining a large number of axial maps, and finds that although there are strong cultural variations in different regions of the world, there are also powerful invariants. The problem is to understand how both cultural variations and invariants can arise from the spatial processes that generate cities. The answer proposed is that sociocultural factors generate the differences by imposing a certain local geometry on the local construction of settlement space, while micro-economic factors, coming increasingly into play as the settlement expands, generate the invariants.

Movement: The strong force

The 'urban grid,' in the sense used in this paper, is the pattern of public space linking the buildings of a settlement, regardless of its degree of geometric regularity. The 'structure' of a grid is the pattern brought to light by expressing the grid as an axial map[8] and analysing it configurationally. A series of recent papers have proposed a strong role for urban grids in creating the living city. The argument centres around the relation between the urban grid and movement. In 'Natural movement',[9] it was shown that the structure of the urban grid has independent and systematic effects on movement patterns, which could be captured by 'integration' analysis of the axial map.[10] In 'Cities as movement economies' it was shown that natural movement – and so ultimately the urban grid itself – impacted land-use patterns by attracting movement-seeking uses such as retail to locations with high natural movement, and sending non-movement-seeking uses such as residence to low natural movement locations.[11] The attracted uses then attracted more movement to the high movement locations, and this in turn attracted further uses, creating a spiral of multiplier effects and resulting in an urban pattern of dense mixed use areas set against a background of more homogeneous, mainly residential development. In 'Centrality as a process',[12] it was then shown that these processes not only responded to well-defined configurational properties of the urban grid, but also initiated changes in it by adapting the 'local grid conditions' in the mixed movement areas in the direction of greater local intensification and 'metric integration' through smaller-scale blocks and more trip-efficient, permeable structures.

Taken together, the three papers describe aspects of a generic mechanism through which human economic and social activity puts its imprint on the spatial form of the city. The papers do not deal with the patterns of activity themselves, but the theory seems to work because, regardless of the nature of activities, their relation to and impact on the urban grid is largely through the way they impact on and are impacted on by movement. Movement emerges as the 'strong force' that holds the whole urban system together, with the fundamental pattern of movement generated by the urban grid itself. The urban grid therefore emerges as a core urban element which, in spite of its static nature, strongly influences the long-term dynamics of the whole urban system. In the light of these results, we can reconceptualise the urban grid as a system of *configurational inequalities* – that is, the differences in integration values in the lines that make up the axial map – which generates a system of *attractional inequalities* – that is, the different loadings of the lines with built form densities and land-use mixes – and note that, in the last analysis, *configuration generates attraction*.

Space-creating mechanisms

The three papers cited describe a process that goes from the spatial configuration of the urban grid to the living city. But what about the grid itself? Is this arbitrary? Would any grid configuration set off the process? The aim of this paper is to try to answer this question. It will be argued that urban grid configurations are far from arbitrary, but in fact are themselves the outcomes of space-creating mechanisms no less generic than the space-to-function mechanisms described in the three cited papers. The argument runs as follows. If we examine a large number of axial maps, we find well-defined invariants as well as obvious differences. What process, we must ask, can produce both? The answer proposed is that the invariants arise from a combination of two things. First, in spite of all their variability, there are certain invariants in the social forces – or more precisely in the relations between social forces – that drive the process of settlement aggregation. Second, there are autonomous spatial laws governing the effects on spatial configuration of the placing of objects such as buildings in space, and these constitute a framework of laws within which the aggregative processes that create settlements take place. The social forces working through the spatial laws create both the differences and the invariants in settlement forms. The link between the two is, again, movement, but whereas the 'space-to-function' mechanism was driven by the effect of spatial configuration on movement, the space-creating mechanism is driven by the influence of movement on space, and so can be considered a 'function-to-space' mechanism.

The concept of spatial 'laws' is critical to this argument, so we must explain what this means. Spatial 'laws', in the sense the term is used here, does not refer to universal human behaviours of the kind claimed, for example, for the theory of 'human territoriality',[13] but to 'if–then' laws that say that if we place an object here or there within a spatial system, then certain predictable consequences follow for the ambient spatial configuration. Such effects are quite independent of human will or intention but can be used by human beings to achieve spatial and indeed social effects. Human beings are bound by these laws in the sense that they form a system of possibilities and limits within which they evolve their spatial strategies. However, human agents decide independently what their strategies should be. Like language, the laws are then, at once, a constraining framework and a system of possibilities to be exploited by individuals.

In fact, it seems likely that human beings already intuitively 'know' these laws (although they cannot make them explicit), and can exploit them as agents to create social effects through spatial behaviours at a very young age. Consider the following true story. A group of people are sitting in armchairs in my daughter's flat. My two-year-old grandson, Freddie, comes into the room with two balloons attached to weights by two pieces of string about two and a half feet long, so that the balloons are at about head height for the sitting people. Looking mischievous, he places the balloons in the centre of the space defined by the armchairs. After a minute or two, thinking Freddie has lost interest, one of the adults moves the balloons from the centre of the space to the edge. Freddie, looking even more mischievous, walks over to the balloons and places them back in the centre of the room. Everyone understands intuitively what is going on, including you. But what is actually happening?

The answer is that by placing an object in the centre of a space we create more obstruction to lines of sight and potential movement than if we place it at the edge. This is the principle of 'centrality' set out in the 'theory of partitioning' in Chapter 8 of *Space is the machine*.[14] If we place a partition midway on a line, it creates more – and more evenly distributed – gain (added distance in summing shortest trips from all points to all others) in the universal distance (the sum of distances from each point to all others) than if we place it peripherally (in which case the depth gain is more unevenly distributed, but is overall less). Because this must apply to lines in all directions, it follows that it will also work for objects placed in space. An object placed centrally in a space will increase universal distance and interrupt inter-visibility more than one placed at the edge. Now it is clear that Freddie not only 'knows' this in the sense that he can make use of this knowledge in behaviour, but it is also clear that he can use this – surely 'theoretical' – knowledge of space to achieve social ends, namely drawing attention to himself and away from the adults engaged in conversation. It is also, of course, clear that we 'know' this about space in the same way as Freddie, but it is also clear, as professionals, that it is unlikely that we were taught this vital principle of space in architecture school or in maths class.

What is proposed here is that spatial laws, driven by social forces, account for exactly and only the spatial invariants of cities.[15] The form of the paper will be to: examine axial maps and develop an account of their invariants as well as their differences; outline and demonstrate the spatial laws in question; apply these to what will be called the 'basic generative process' by which urban-type spatial systems arise; and develop a theory of how the impact of the spatial laws on evolving settlements is driven by two kinds of social forces, which can be broadly termed the sociocultural and the micro-economic. It is proposed that culture is a variable and puts its imprint mainly on the local texturing of space, generating its characteristic differences, whereas micro-economics is a constant and puts its imprint mainly on the emerging global structure of the settlement in a more or less invariant way. The reason one works locally and the other globally is due to the ways in which each uses the same spatial laws to generate or restrain potential movement in the system.

This is why we find, in axial maps, both differences in local texture, and invariants in the global patterning. The combination of the spatial laws and the dual processes explains why axial maps read as a set of similarities and differences. The paper concludes with a discussion of the relation between socio-economic and spatial laws, suggesting that although the creation of the space of the city is driven by socio-economic processes, it is not shaped exclusively by them. Equally fundamental in shaping city space are autonomous spatial laws that generate more or less equifinal outcomes from varying processes.[16]

Differences and invariants in axial maps

First, let us consider some axial maps. By far the most obvious differences between them are geometrical. On reflection, that is all they could be. Axial maps are no more than sets of lines of different lengths with different angles of intersection and different degrees and kinds of intersection (for example, a line can either pass through another or stop on it). Axial maps from different parts of the world tend to differ in all these properties. Figures 17.1–17.4 show four fairly characteristic axial maps from

Figure 17.1 Atlanta axial map. Credit: Mark David Major.

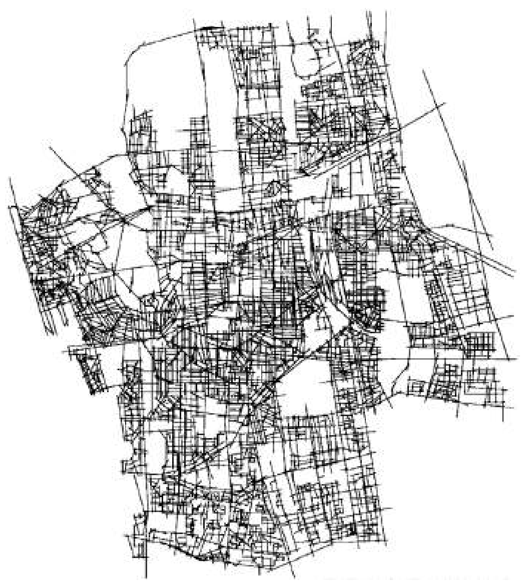


Figure 17.2 The Hague axial map. Credit: Laurie Neale.

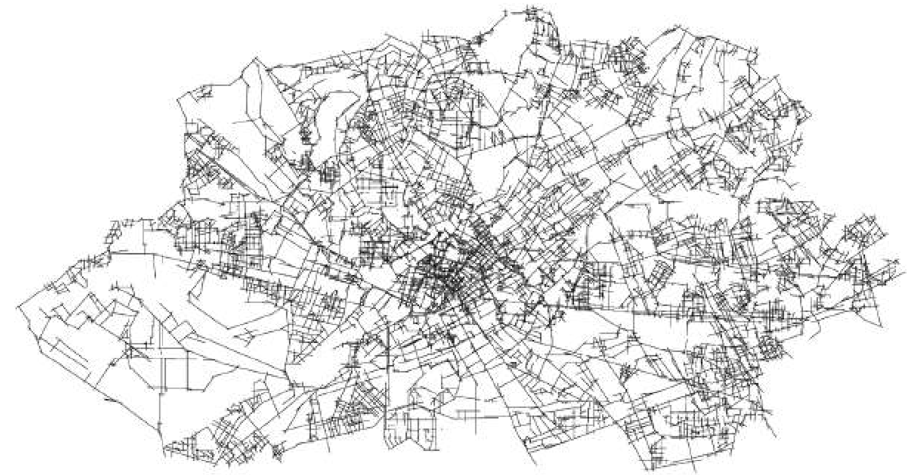


Figure 17.3 Manchester axial map. Credit: Polly Fong.

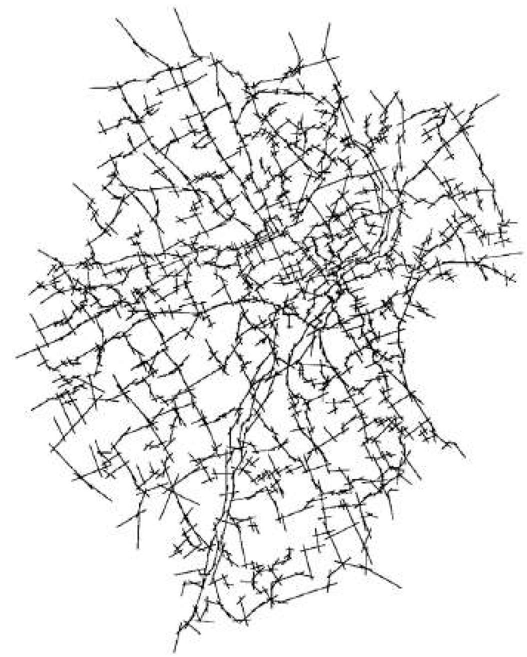


Figure 17.4 Hamedan axial map. Credit: Kayvan Karimi.

different parts of the world arranged from the most to the least 'geometric': Atlanta (USA), The Hague (Holland), Manchester (UK) and Hamedan (Iran). It is easy to see that the impression of 'more to less geometric' arises because the axial maps differ substantially on their basic properties. Each has its own distinctive range of line lengths and angles of incidence, and its distinctive intersection characteristics. For example, if we use the patterns of intersection, we find that in Atlanta, the tendency for lines to pass through each other (rather than to end on other lines) is very marked at all levels. In The Hague, this is found locally but much less so at the global level. In Manchester, this is hardly found globally, and what there is locally is much more broken up than in The Hague. In Hamedan, it hardly exists either at the global or local levels, except in the central public areas of the town.

Differences in the range of line lengths and angles of incidence seem to follow the intersection differences. Atlanta has a number of very long lines approximating the radius of the system, and long lines can be found in most parts. At the same time, large areas of the grid maintain a strict right-angle intersection with a north–south orientation, although with a striking offset grid in the historic centre. In The Hague, the longest lines tend to be less than the radius of the system and, in general, long lines are peripheral to discrete local groups of lines. A less strong, right-angle discipline is maintained, there is greater variety in the orientation and long lines, especially radials, tend to intersect with other lines at very obtuse angles. In Manchester, the long lines are nearly all radials well below the radius of the system. The tendency for the long radials to have near-straight continuations is even stronger, and the local right-angle discipline is even looser. In Hamedan, the longest lines are only a fraction of the radius of the system and tend to be found towards the periphery of the system. Even so, there is a clear radial structure formed by lines of the second length rank, and intersecting with greater angular change than in the other cases. Locally, we find a whole range of angles of incidence, including near-right-angle connections, but in most cases one line tends to stop on another.

These geometric differences are also consistently reflected in syntactic differences.

Table 17.1 shows the syntactic average for 58 cities taken from four parts of the world. Each regional group of cities, in spite of differences within the subsamples, has its own characteristic set of syntactic parameters.

What is the reason for these geometrical and syntactic differences? Why should lines in Iranian cities be, on average, markedly shorter than lines in, say, English cities, or why should European cities have a degree of geometric organisation somewhere between UK and American cities, or Arab cities be less intelligible than European cities? On the face of it, the differences seem to be expressions of what we might call 'spatial culture'. For example, in cities in the Arab World, the spectrum between public and private spaces is often quite different from that in European cities. In historic European cities, we find that local areas are, for the most part, easily permeable to strangers, with public spaces in locally central areas easily accessible by strong lines from the edge of the area. At the same time,

Table 17.1 The syntactic average for 58 cities taken from four parts of the world

	Cities	Avg. lines	Conn	Local int	Global int	Intel
USA	12	5,420	5.835	2.956	1.610	0.559
Euro	15	5,030	4.609	2.254	0.918	0.266
UK	13	4,440	3.713	2.148	0.720	0.232
Arab	18	840	2.975	1.619	0.650	0.160

Conn = connectivity; Local int = local integration; Global int = global integration; Intel = intelligibility.

fronts of dwellings are strongly developed as facades and interface directly with the street both in terms of visibility and movement. In many Arab cities, strangers tend to be guided much more to certain public areas in the town, and access to local areas is rendered much more forbidding by the more complex axial structure. At the same time, dwelling facades are much less developed, and the interface with the street tends to be much less direct both for visibility and for movement. The differences in the geometry of the axial maps seem to be a natural expression of these differences. Even in the case of American cities, where one of the main factors in creating the more uniform American grid is thought to be the need to parcel up land as quickly and easily as possible to facilitate economic development, we note that the grid was prior to economic development and should therefore be seen as a 'spatial cultural' decision to create and use space in a certain way.

However, in spite of these differences, there are also powerful invariants in axial maps that seem to go across cultures and even across scales of settlement. One of the most striking is the statistical distribution of line lengths. Although we find great variations in the average and range of line lengths, we invariably find:

- that the axial maps of cities are made up of a small number of long lines and a large number of short lines;
- that this becomes more the case as cities become larger; and
- that in general the distribution of line lengths in cities approximates a logarithmic distribution.

Figure 17.5 shows the four cities of Figures 17.1–17.4 with the distribution of line lengths on the left and the logged distribution on the right.[17]

In practical terms, this means that if, for example, we divide the range of line lengths into 10, we find that in Atlanta 92.7 per cent of lines are in the decile of shortest lines and only 2 per cent in the eight longest. In The Hague, the figures are 84.8 and 5 per cent, and in Manchester 85.9 per cent and less than 3 per cent. In the much smaller case of Hamedan, we find that 90 per cent of the lines are in the four shortest deciles and only 2 per cent in the five longest. Looking more widely, we

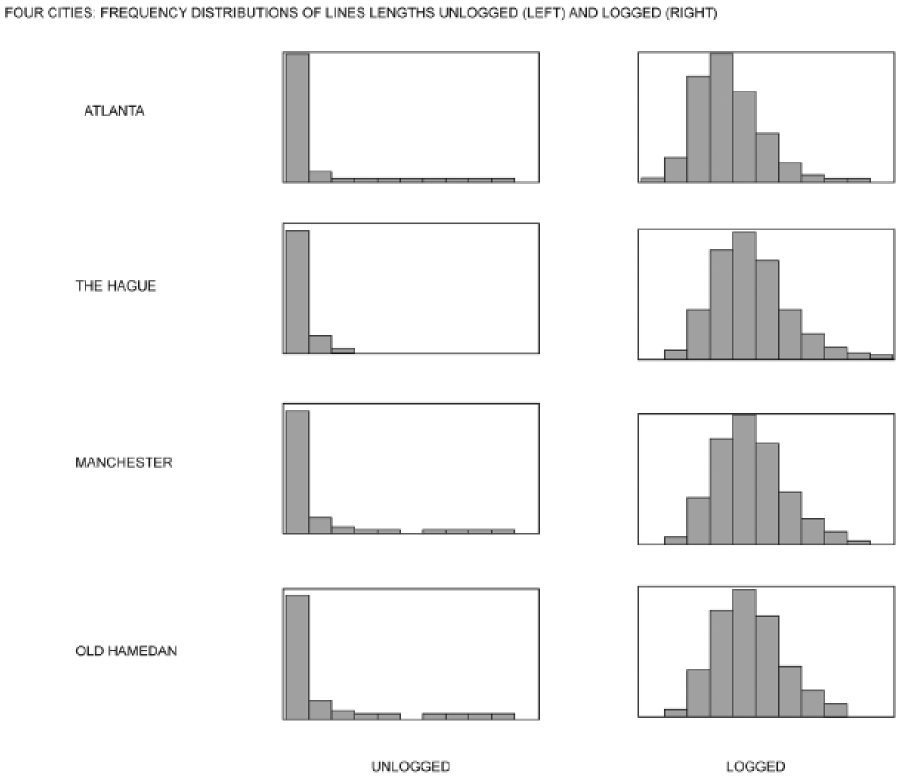


Figure 17.5 The four cities of Figures 17.1–17.4 with the distribution of line lengths on the left and the logged distribution on the right.

find that in London (15,919 lines), 93.3 per cent of lines are in the shortest decile and less than 1 per cent in the top eight deciles. In Amsterdam (7,996 lines) – on the face of it a more griddy city – the figure is 95.8 per cent in the shortest decile and again less than 1 per cent in the top eight. In Santiago (29,808 lines), an even more grid-like structure, the figure is 94.7 per cent, again with 1 per cent in the top eight, while in Chicago (30,469 lines), a city we think of as wholly grid-like, the figure is 97.6 per cent with only 0.6 per cent in the top eight. It is not always quite so high in large cities. In Athens (23,519 lines), for example, the figure is 86 per cent, with 2.3 per cent in the top eight. However, even in the strange pre-Columbian city of Teotihuacan the figure is 85 per cent.

If we look at a smaller system, we find the same tendency, although less marked. In Venice (2,556 lines), for example, the figure is 76.3 per cent with 4 per cent in the top eight; in Hamedan (1,971 lines) in Iran, where lines are on average shorter than in Western cities, 71.7 per cent are in the shortest decile and 8.3 per cent in the top eight. In the English cities of Nottingham, Bristol and York, the figures for the shortest decile are 78, 63 and 55 per cent, respectively. Even in

much smaller systems, we can find a strong tendency in this direction. If we take Old Paranoá, the informal settlement built by the workers who constructed the dam in Brasília, we find that 32 per cent of the lines are in the shortest decile and 68 per cent in the shortest two.[18] In the southern French town of Apt, 41 per cent are in the shortest decile and 59 per cent in the shortest two, while in Serowe (a self-generated settlement in southern Africa), 32 per cent are in the shortest decile and 68 per cent in the shortest two. Even in a small area within London we find 24 per cent in the shortest decile and 53 per cent in the shortest two.

As settlements grow, then, the proportion of lines that are long relative to the mean for the settlement becomes smaller, but the lines themselves get longer. This seems to be invariant across all cultures in spite of the strong geometric differences we have noted. A plot of the log of the number of lines against the proportion in the shortest decile for 20 settlements from small to very large shows an r^2 of 0.802 ($p = 0.0001$). This also applies to different-sized chunks of the same city. If we plot the percentage of lines in the shortest decile against the number of axial lines for four different-sized cut-outs from the London axial map, the graph shows an r^2 of 0.923, $p = 0.0391$. However, even the smallest cut-out – the City of London with only 565 lines (as opposed to 15,919 for the largest system) – has 70 per cent of lines in the shortest decile, and approximates a logarithmic distribution.

Why then are line lengths distributed in this way and, in particular, what is the role of the small number of long lines? A useful clue comes from looking at their spatial distribution. If we take the lines in the longest quintile of the range and make them the darkest lines in the axial map, we find a marked tendency for the longest lines to be centre-to-edge lines starting at some distance from the original centre. Figure 17.6 shows the pattern for London and Figure 17.7 for Athens. The second rank of lines, however, shows a different pattern in each case. In London, the second-rank lines form a continuous and relatively dense network penetrating

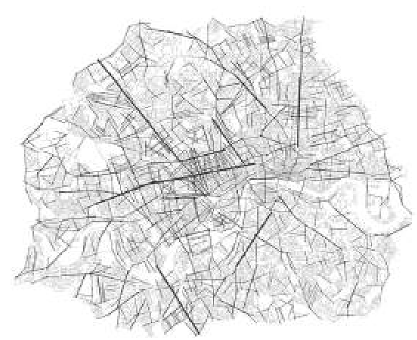


Figure 17.6 London: line length. Credit: Chang Hua Yoo.

Figure 17.7 Athens: line length. Credit: Valentina Karvounzi.

most parts of the grid. In Athens, the second-rank lines pick out discrete, grid-like areas with relatively poor connections between them. If we were to look at, say, Baltimore, the second-rank lines tend to be linked directly to the first rank of lines, forming a tree-like distribution in the system as a whole. These patterns suggest that the first rank of lines reflects generic properties of city growth while the second rank indicates differences in the relation of global to local.

This hint of global invariants and local differences is reinforced if we look at the syntactic analysis of the axial maps. If we take the four cities shown in Figures 17.1–17.4 and analyse them for radius *n* integration (for example, Figure 17.8), we find in each case that in spite of the geometric differences a certain kind of structure is adumbrated: each city has an 'integration core' – the patterns formed by the darker lines – which links a grid-like pattern of lines at the heart of the city almost to the edge in all directions either by way of quasi-radial lines or extended orthogonal lines, in some cases reaching the edge line but in others falling short. Within the interstices formed by this integration core, lighter areas are found, often with a darkish line as a local focus. In other words, in spite of the geometric differences, each city has, when seen as a system of configurational inequalities, a certain similarity of structure. This is the pattern we call the 'deformed wheel': a hub, spokes in all main directions, sometimes a partial rim of major lines, with less integrated, usually more residential, areas in the interstices formed by the wheel. This generic pattern was first identified as a deep structure common to many small towns, seeming to occur in spite of topographic differences.[19] It was also found as a local area structure in London, where named areas such as Soho or Barnsbury were typically identified as deformed wheel forms, with the London 'supergrid' (the main radials and their lateral links) forming the rim of the wheel.

As a global pattern, the deformed wheel holds up remarkably well in larger cities, for example, in London, Athens and Baltimore. The pattern is even found in very different kinds of cities. If we look at Venice (Figure 17.9) without the canals, for example, we find that in spite of its very idiosyncratic history – having grown together from several islands rather than from a single origin – we still find a very marked 'deformed wheel' pattern, even though the wheel is much less easily recognisable than in most cases. Or looking at Tokyo, which is by far the largest system ever analysed, we find a remarkable and even more complex version of the wheel pattern with several layers of rim which, with the sinuous radials, produce a quasi-grid that covers a large part of the system. Even the strange pre-Columbian city of Teotihuacan shows at least a partial realisation of the deformed wheel pattern. Again, this near invariant of cities is found in spite of the substantial differences in syntactic values that were shown in Table 17.1.

In addition to these space invariants, we also find that if we look at settlements in terms of the size and shape of blocks, then we find, if not invariants, then at least a set of pervasive tendencies, once again set against a background of substantial geometric differences by region (which we may therefore expect to have a cultural origin of some kind). These can be seen fairly easily in the axial

Figure 17.8 Atlanta axial map global (radius *n*) integration. Credit: Mark David Major.

Figure 17.9 Venice axial map global (radius *n*) integration. Credit: Erica Calogero.

maps of Figures 17.1–17.4, and even in the analysed axial maps, but is perhaps easier to see in black-on-white figure ground maps of a Turkish city analysed by Sema Kubat, shown in Figure 17.10. The most obvious 'near invariant' is an underlying tendency for blocks to be smaller and more convex at or near the centre and larger and less convex towards the edges.[20] However, if we relate block size and shape to the patterns shown by integration analysis of the axial map, we find a subtler pattern. The lines forming the spokes of the deformed wheel tend to be lined with larger than average (for the settlement) blocks for most of their length, but smaller than average blocks in the centre (the hub of the wheel).[21] In contrast, the areas interstitial to the core tend to have block sizes between these two extremes. In other words, the distribution of block sizes seems to reflect the distinction between global and local structure. This is to some extent the case in all the settlements shown so far.

Socio-cultural relativities and economic universals

We are faced then with a puzzle. The processes that generate the axial maps and block maps of cities seem at the same time to produce variants, in the form of systematic differences in settlement geometry and syntax from one region to

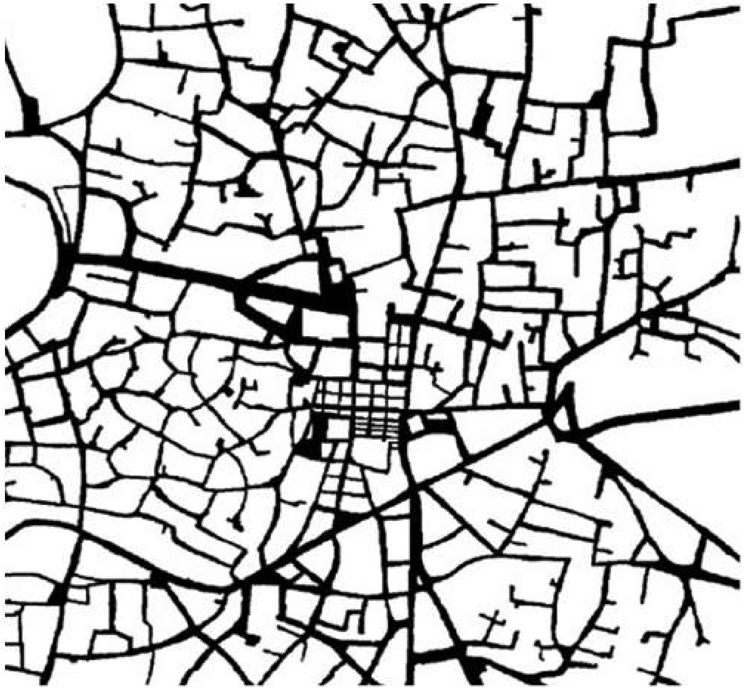


Figure 17.10 Figure-ground map of Ancient Konya, Turkey. Credit: Ayse Sema Kubat.

another, and also invariants. What kind of process can produce both? It seems highly unlikely that these dual patterns are in any sense 'designed in', although of course they may be in some cases. However, the fact that most settlements evolve over long periods compels us to the view that the patterns arise from a largely 'distributed' or 'bottom-up' process, that is, from multiple interventions by many agents over time. Even if single agencies are involved, then even so the fact that settlements evolve over such long periods implies that the process of settlement generation must be regarded as an essentially distributed one. What kind of distributed process, then, can produce such dual emergent phenomena?

Let us first note an important difference between variants and invariants: the variants tend to be local and the invariants global. Now consider a case where a city has grown under the influence of at least two different cultures: Nicosia. Figure 17.11 is an analysed axial map of the historical core of Nicosia in Cyprus (a city sadly now divided). The northeast quarter is a historic Turkish area, the southeast a historic Greek area. The differences in the texture of the grid are marked, with the two areas having quite different geometries and different emergent topologies: the Greek area has longer lines, more lines passing through each other, a different pattern of angle of incidence, and, as a result, much more local and global integration (and a better relation between the two) than the Turkish area. Since these differences reflect typical differences found between systems in Europe and the Islamic world, it is reasonable to regard these as socio-cultural differences in the basic geometry of space. However, when we analyse the area as a whole we find a typical deformed wheel pattern has somehow arisen over and above these geometrical differences, even though the differences between the Greek and Turkish areas show up strongly as differences in the degree of integration.

We thus see what appear to be two processes operating in parallel: one a local process generating differences in local grid patterns and apparently reflecting differences in spatial culture in some way; and the other a global process generating a single overriding structure that seems to reflect a more generic or universal process of some kind. A clue to this comes from the simple fact that the less integrated areas generated by the local process are largely residential, and it would be natural to think of these as the primary distributed loci of sociocultural identities, it being

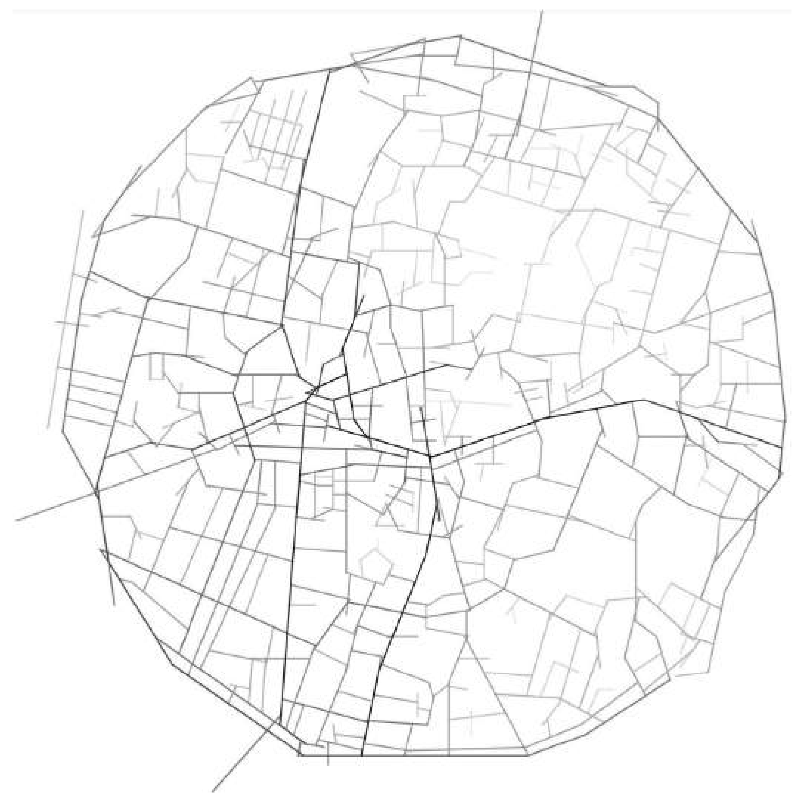


Figure 17.11 Nicosia axial map global (radius *n*) integration. Credit: Konstantinos Kypris.

through domestic space and its environs (including local religious and cultural buildings) that culture is most strongly reproduced through the spatiality of everyday life. A second clue comes from the fact that when we analyse settlements syntactically, it is the micro-economic activity of markets, exchange and trading that is most strongly associated with the 'Integration core', religious and civic buildings being much more variably located.[22] In this, of course, the integration core of public space also reflects the spatiality of everyday life, but in this case it tends both to the global, because micro-economic activity in its nature will seek to extend rather than confine itself, and also to be culturally nonspecific, in that it is in these activities, and therefore these spaces, that people mix and cultural differences are backgrounded.

This suggests a natural explanation for the dual production of variants and invariants in urban grids. On the one hand, a residential process driven by socio-cultural forces puts its imprint on local space by specifying its geometry and generates a distinctive pattern of local differences, because culture is spatially specific. On the other, a public space process driven by micro-economic activity generates a globalising pattern of space that tends to be similar everywhere because micro-economic activity is a spatial universal. This is the critical difference between the two aspects of the settlement-creating process: the socio-cultural component is idiosyncratic and local while the micro-economic component is universal and global. It is this that creates the underlying pattern of differences and invariants that we find everywhere in settlement forms.

This is the key conjecture of this paper: that the processes that generate settlement forms are essentially dual, and through this duality generate the invariant pattern of local differences and global similarities that characterise settlement forms. The question then arises: why should sociocultural life generate one kind of spatial pattern and micro-economic life another? The answer, it will be proposed, lies in the fact that the relation between micro-economic activity and space, like the relation between culture and space, is largely mediated by movement, but with regard to micro-economics in a universal and global way, and with regard to culture in a local and specific way. In what follows we will therefore look at spatial and movement aspects of both socio-cultural and small-scale economic processes and how they affect each other as a settlement grows.

The basic generative process

We can begin by noting that there is also a set of low-level invariants, or near-invariants, in urban space, which are so commonplace as to be rarely remarked on, but which are the very foundation of what a settlement is. These are:

- that most spaces are linear, defined by the entrances of buildings or groups of buildings on both sides;
- that buildings are clumped together to form discrete islands;
- so that the linear spaces surroundings the islands form intersecting rings and create an overall system of continuous space (a 'street pattern' of some kind); and
- that this is a highly nondendritic configuration, that is a pattern that is everywhere ringy rather than tree-like.

The simplest process for generating spatial configurations with these properties has been familiar since the earliest days of space syntax: the restricted random 'beady ring' process that generates small ring street settlements of a kind found in many parts of the world.[23] The process starts with a dyad composed of a cell (representing a notional building) and a piece of open space linked by an entrance so that those inside can come and go into the outside world. These dyads aggregate randomly apart from two restrictions: that each open cell must join full-facewise onto one already in the system (joins of closed cells arise only randomly); and no vertex joins for closed cells are allowed (people do not build corner to corner).[24]

The pattern on the left of Figure 17.12 is a typical product of such a process. A 'beady ring'-type pattern is produced on the way, but this is not our main concern here. The overall pattern is that a system of outward-facing islands of built forms, varying in size and creating more or less linear spaces forming intersecting rings, has emerged from the process. No one designed this. It has emerged by a process that finds a pathway of emergence by which a global pattern appears from the actions of local agents. A key element of the urban system has thus emerged in the form of a continuous system of open space, permitting interaccessibility from each part of the settlement to all others.

The pattern thus has enough of the key topological settlement-like properties (although it lacks their geometric properties – but see below) for us to think of it – by Ockham's Razor perhaps – as the 'basic generative process' for spatial patterns of a generically urban kind. But it does not yet look at all like a real settlement. What is missing? It cannot be just the over-regularity due to the fact that the process has been generated on a regular underlying grid. The fault seems to lie mainly in the geometry of both its block structure and its line or axial structure: blocks are insufficiently compact and lines are insufficiently varied in length. So let us look at two real settlements that seem to have grown by something like this process and see what they have in addition. Figure 17.13 is the old self-generated settlement of Paranoá, developed from the encampment of the workers who built the dam for the lake behind Brasília.[25] Figure 17.14 is the settlement of Serowe in southwest Africa in which the built elements are actually compounds. On the right are radius *n* integration maps of each, and the 'synergy' scattergram plotting the correlation between local and global integration. On the right of Figure 17.12 is the same analysis of the computer-generated pattern.

Two points are of particular interest. First, something like the deformed wheel 'integration core' exists in both real cases (and in the case of Paranoá cannot

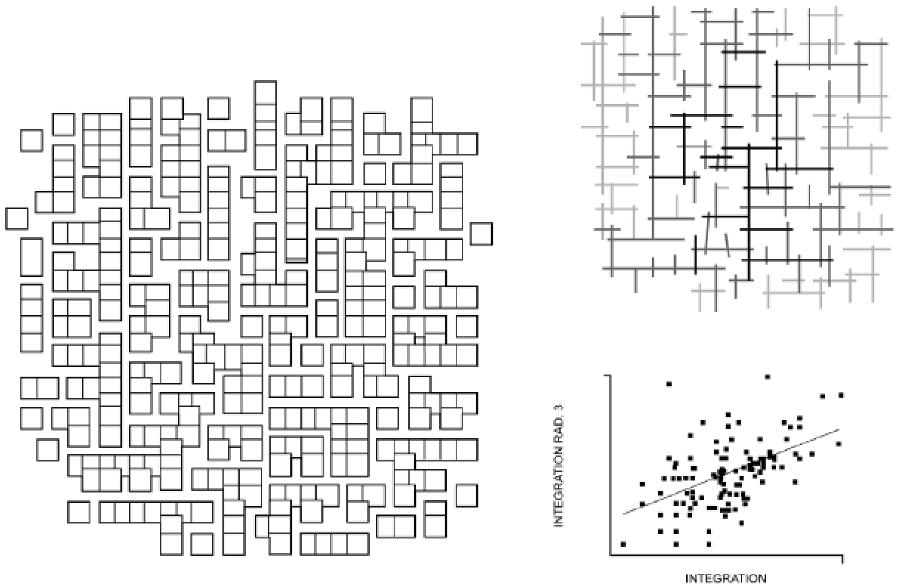


Figure 17.12 Computer-generated surface showing emergent block structure and analysis of emergent space pattern.

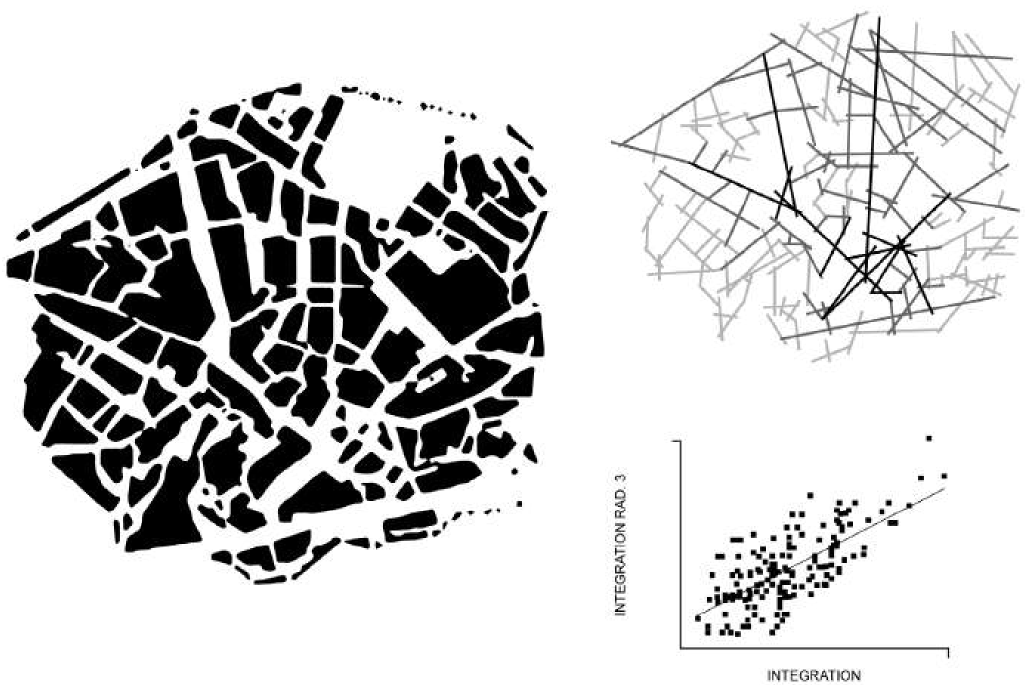


Figure 17.13 Space structure and analysis of Paranoá, the settlement that grew from the encampment of the builders of the dam in Brasília. Credit: Frederico de Holanda.

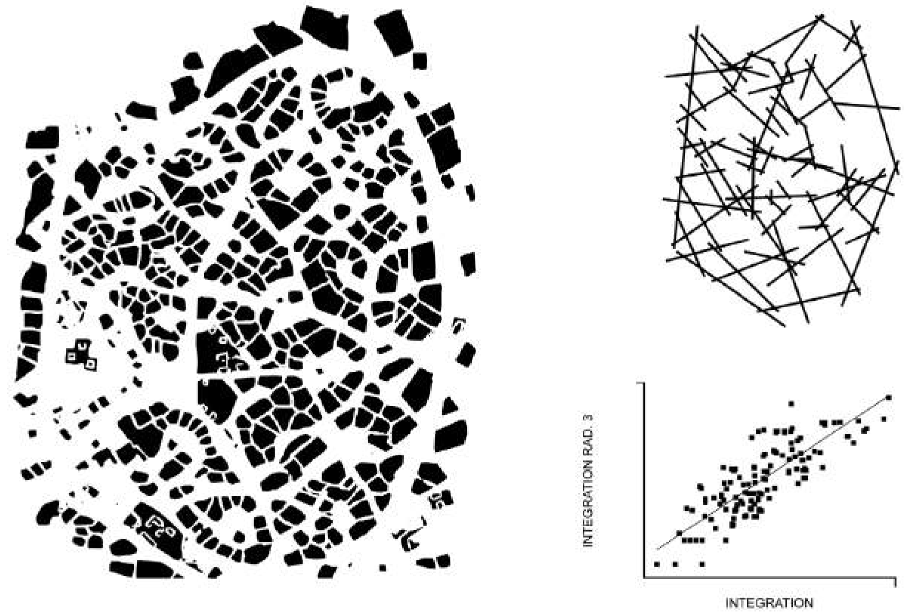


Figure 17.14 Serowe with radius *n* axial analysis and synergy scattergram. Credit: Glenn Mills.

be explained in terms of existing routes in the direction of other settlements, since there were none except to the south). Second, when we look at the 'synergy' scattergrams, we find that the r^2 between local (radius 3) and global (radius n) integration is much better than in the generated case in spite of the fact that it lacks the discipline of an underlying grid. In other words, Paranoá and Serowe both display a relation between local and global structure that needs to be explained.

Experiments with random lines

We can explore these differences further by experimentation. We first construct a more or less random rectilinear grid made up of lines that vary in length by only a small amount, on average about half the diameter of the overall 'settlement'. The scattergram gives an r^2 between connectivity and integration of over 0.8[26,27]. We then retain the same mean and range of line length but grow the system to twice its size. Its diameter is now about three times the mean line length. The 'intelligibility' r^2 falls to 0.5. We do the same again, increasing the size of the system until its diameter is about four times the mean line length. The r^2 falls to below 0.3, as in Figure 17.15.

It is not difficult to work out what is happening. If integration analysis is carried out on a system with uniform elements much smaller than the system

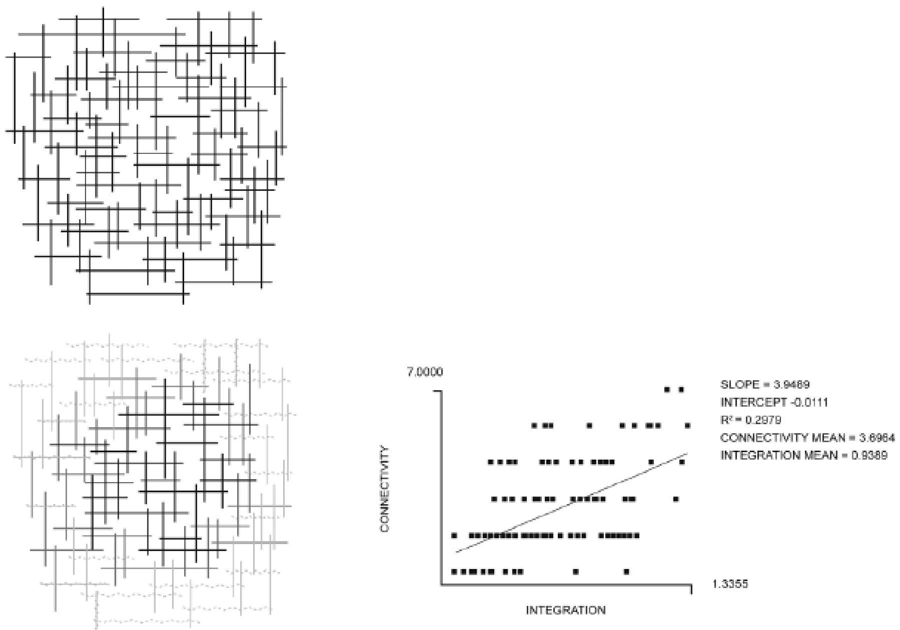


Figure 17.15 Random system with radius *n* axial analysis and intelligibility scattergram.

itself – say a tessellation of square cells – then integration will focus on the geometric centre and fall off towards the edges. As soon as you specify a system with more or less similar dimensions, in this case similar line lengths, then the same must happen. As the system becomes larger, integration will increasingly concentrate in the centre. The consequences for the 'intelligibility' relation between line connectivity (which is closely related to length) and integration are that relatively longer and therefore better-connected lines will be randomly distributed through the system, while integration will be concentrated in the centre. The more this happens, the less the two will correlate and the more the local properties of the system give a poor guide to the global properties – hence unintelligibility.

If we then take four lines near the centre and extend them to a length of about 0.75 of the diameter of the system, the effect on both the integration core and the scattergram is immediate and dramatic. The core, not surprisingly, begins to go from centre to edge and the scattergram improves from below 0.3 to above 0.6. However, the scatter is highly non-urban, in that the four new lines are quite distinct from the rest of the system. But in Figure 17.16, an r^2 of .86 is achieved with a pattern of lines that links laterally at the edges as well as from centre to edge: the characteristic deformed wheel structure. In this axial map, 47 per cent of lines are in the shortest decile and a further 29 per cent in the next shortest, almost identical to Paranoá, where the respective figures are 52 and 25 per cent.

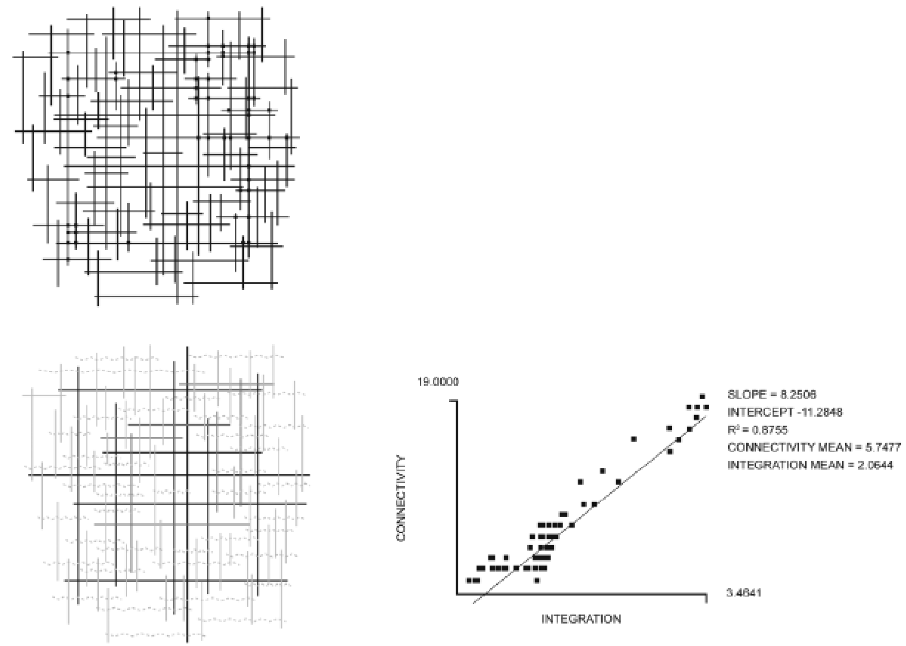


Figure 17.16 Adapted random system with radius *n* axial analysis and intelligibility scattergram.

This suggests that the essential function of the longer lines against the background of shorter lines is, as we might expect, to give some kind of global structure to the overall pattern, with the local structure fitted into its interstices. However, two further points must be added. First, we also find that the pattern of long to short lines is critical not just to the global structure but also to the relation between the local and global structure. This suggests that the long to short distribution is pervasive at all levels of the settlement and its growth, and therefore needs to be understood as an outcome of a growth process rather than as one of imposition of a global structure. In other words, we need to understand how the required distribution of line lengths can be produced at every stage of an aggregative process of settlement growth.

Regularities in the configurational effects of placing objects

How then can we modify the 'basic generative process' to create these outcomes at every level, so that the growth process will tend not only to create a pervasively lognormal distribution at every level but at the same time also to generate an intelligible and synergic system with a deformed wheel-type structure? The answer proposed is that it is here that spatial laws intervene, driven by the dual sociocultural and micro-economic forces imposing on space their different requirements for potential movement.

The laws in question govern the effects on spatial configuration of the placing objects (such as buildings) in space. The laws initially govern the degree of metric integration in the system, measured as the universal distance[28] from each cell in the complex to all others (as opposed to a specific distance that measures distance from one cell to one other). The mean universal distance in a complex is thus isomorphic to the mean length of the trip by shortest paths within the complex. It is through their effect on mean trip lengths that these laws are activated and govern the evolution of the urban object.

The laws are essentially clarifications, simplifications and fuller demonstrations of the 'principles of partitioning'.[29] In chapter 8 of *Space is the machine*, it was shown that every time a partition is placed in a system, it has a predictable effect on 'universal distance' within that system. In that text, four partitioning principles were proposed for the minimising or maximising of 'depth gain' in a system, depth gain being the increase in universal distance due to the placing of a partition.[30] The principles were: *centrality* – partitioning a line in its centre creates more depth gain than partitioning it eccentrically; *extension* – partitioning a longer line creates more depth gain than partitioning a shorter line; *contiguity* – making partitions contiguous increases depth gain more than making them discrete; and *linearity* – arranging contiguous partitions linearly increases depth gain more than coiling them up, as, for example, in a room.

In what follows it will be proposed that these four principles[31] can be reduced to two laws, one dealing with the relations of spaces and the other with the relations of objects. Before we introduce these laws, however, we will show how these partitioning regularities can be interpreted for cellular aggregates. The basic notion we work with is that of a pair of cells (or boundaries) forming the two ends of a line, and a third cell which we wish to place between them. The method for calculating the gain in universal distance is as in Figure 17.17. Consider a line of $n+1$ cells with an object placed somewhere along it, leaving n cells in some distribution on the two sides of the cell with at least one cell on each side. A deviation, d, will be the unit distance around the object that must be added to straight line movement to go from any cell to any other on the other side of the object. D will be the sum of ds that are needed to go from all cells to all others, or the total added to the universal distance on that line by the placing of an object.

If the object is square, and its sides are the same size as the unit of distance for measuring along the line, as in the top case in Figure 17.17, then d will always be 2 units of distance. Here we refer to the 2-unit deviation as a single d. Note that if an object with, say, shape 3×1 is placed on the line lengthwise, then, as in the middle case in Figure 17.17, d for negotiating that object will always be 2 units regardless of the length of the unit, because the trip between the two deviation units is parallel to the original line. If, however, the 3×1 object is placed orthogonal to the line of movement (see the bottom case in Figure 17.17), then a further two

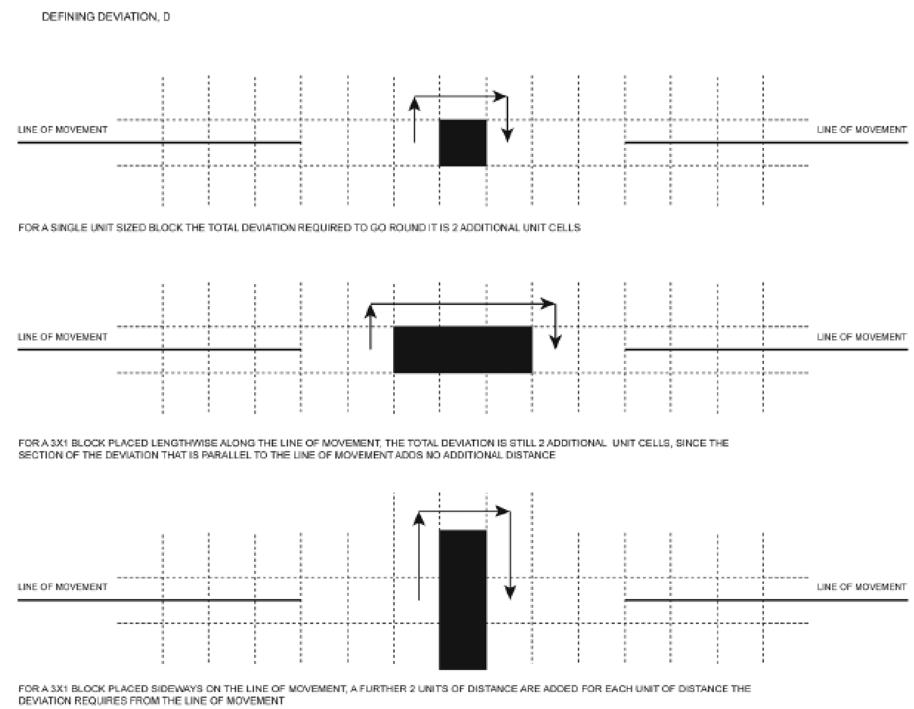


Figure 17.17 Partitioning example.

units of distance, that is one further d, will be added for every parallel line blocked by the object.

Figure 17.18 illustrates the principle of centrality: if we want to place a cell (the light cell top left) between two existing cells (dark), does it make a difference where we place it? The answer (mid-left) is that the more peripherally we place it, the less the increase in universal distance, and the more centrally we place it, the greater the increase. It follows (bottom left) that if we place cells evenly along lines, the increase in universal distance is greater than if we make some gaps large and others small. It also follows (mid-right) that an object placed in the centre of a space will increase universal distance more than one placed towards the edge (because the effect on two dimensions will be the sum of linear effects). The principle of extension also follows: if we place a block on a longer line, it increases universal distance more than if we place it on a shorter line. Figure 17.19 illustrates the principle of contiguity: cells joined contiguously increase universal distance more than those placed discretely. Finally, Figure 17.20 illustrates the principle of linearity: contiguous cells arranged linearly increase universal distance more than if they are placed compactly. It should be noted that these principles interact.

The impact of these laws on grids can be explored by constructing experimental grids made up of metrically uniform cells (they can be as small as

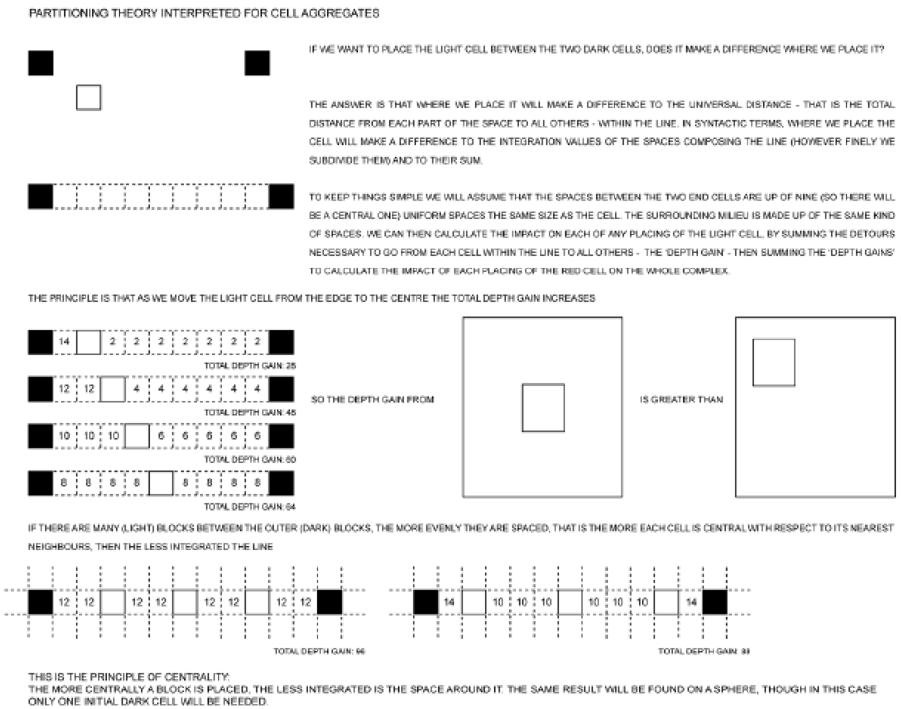


Figure 17.18 Partitioning: example of contiguity.

we like, as long as they are uniform), and calculating the mean universal distance, or mean trip length, for each. Figure 17.21 sets out a series of experiments with grids each with 301 metrically uniform cells. The cells are circular in order to avoid the effect of corner joins. Each grid thus has the same number of metric cells and therefore the same number of distance elements. Differences between grids are therefore purely to do with the rearrangement of the cells into different configurations. In some cases, the rearrangement has left one cell that cannot be located in the grid. In each of these cases, the cell has been added to the same position in the grid, namely the intersection of the third column (counting from the left) and the third row (counting from the top). Experiments with the sensitivity of the grid to the addition on one overlaid cell show that an additional cell overlaid in the centre of the uniform regular grid reduces the mean universal distance by 0.1 per cent (it will of course slightly increase the total depth since there is an additional cell), while overlaying it on a corner cell increases it by 0.2 per cent. These differences are then one or two orders of magnitude less than the effects of configurational changes below, and so can be discounted.

Using the regular uniform grid (Grid A) as the benchmark, we can then vary the configurations of grids to illustrate the effect of the four principles. In Grids B

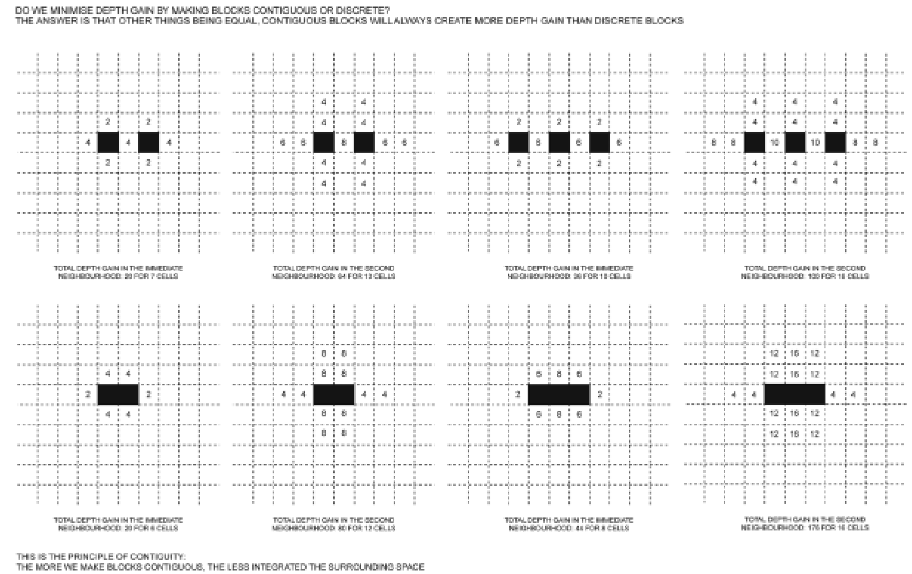


Figure 17.19 Partitioning: example of centrality.

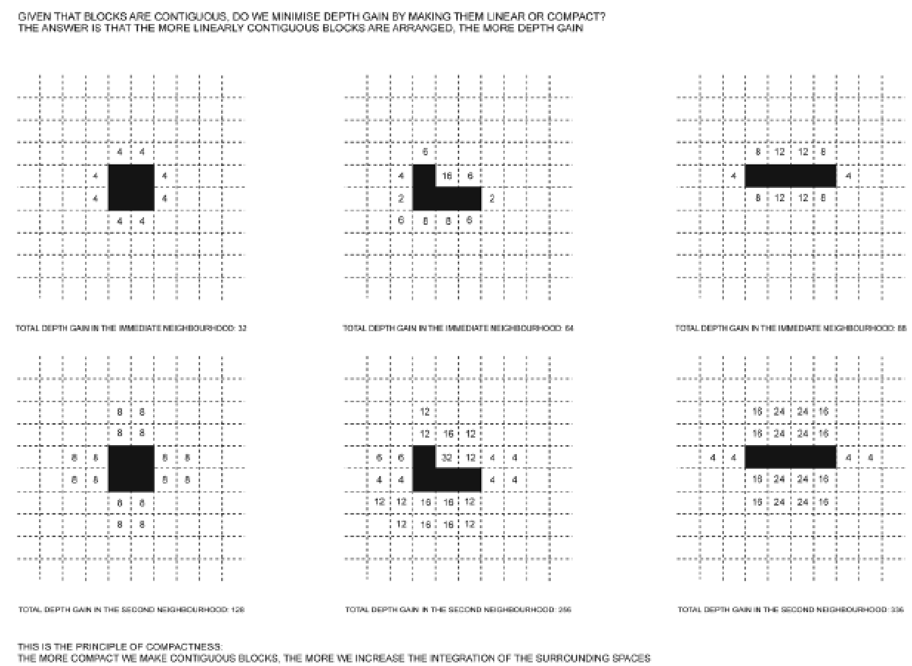


Figure 17.20 Partitioning: example of compactness.

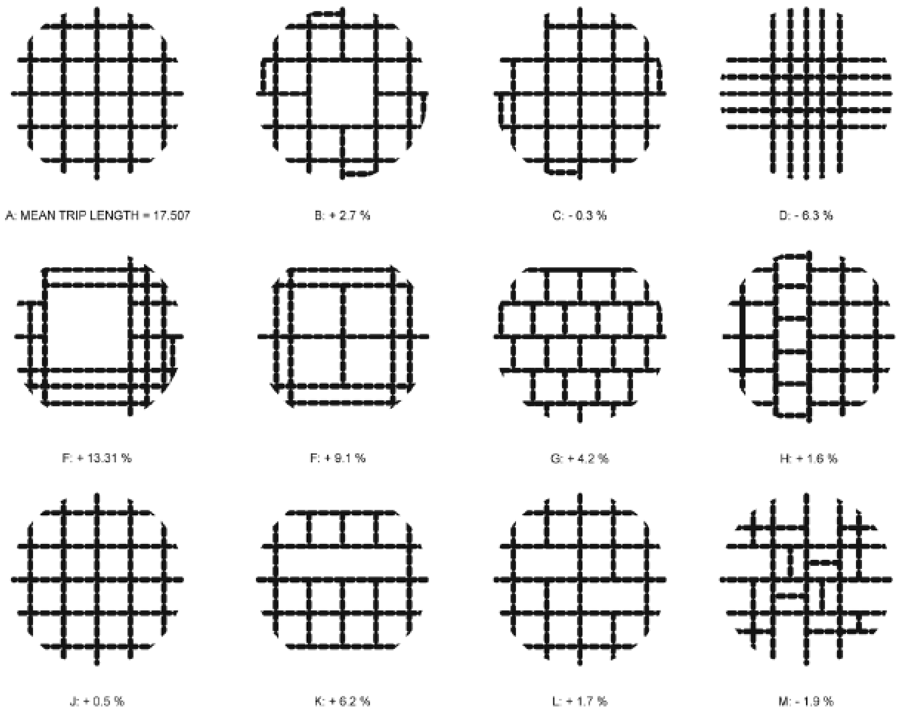


Figure 17.21 Experimental grids compared to regular grid with regard to mean trip length.

and C, for example, we illustrate the 'centrality' principle by placing a block initially in the centre and then in the corner, while standardising the layout of the cells displaced from the centre. Placing the block in the centre increases the universal distance of the grid by 2.6 per cent, while placing it in the corner reduces it by 0.3 per cent. In Grid D, we take this further by reducing the scale of blocks in the centre at the cost of increasing them at the edge (a common form in the centre of towns, as noted in 'Centrality as a process').[32] The mean universal distance is reduced by 6.3 per cent. If we do the opposite and make the centre block as large as possible, and place the small blocks at the edges (the number of small blocks remains the same, as in Grid E), we increase the mean universal distance by 13.9 per cent, making a total difference between Grids D and E of just under 20 per cent. In Grid F, we take Grid E and create a cross link through the centre. The effect is to increase mean universal distance by 9.1 per cent compared to Grid A, but to reduce it by nearly 5 per cent compared to Grid E.

We then illustrate the principle of extension. In Grid G we displace each vertical segment of cells between grid intersections one cell to the right and then to the left on alternate rows. We, thus, shorten all internal vertical lines with more or less neutral effects on block sizes. The effect is to increase mean universal distance by 4.2 per cent. In Grid H, we break all horizontal lines close to the centre vertical, creating a pair of lines of fairly equal length at each level. The increase in mean universal distance is 1.6 per cent. However, when we break the horizontal line near the edge vertical in Grid J, thus keeping some lines as long as possible at the expense of others becoming much shorter, the mean universal distance increases by only 0.5 per cent, three times less than with a more central break in the lines.

The principle of compactness is illustrated in Grid K by converting the square central block of Grid B into a linear block of equal area. The effect is to increase the universal distance by 6.2 per cent compared to 2.6 per cent for the square block. We then illustrate the principle of contiguity by splitting the linear block into two in Grid L. The increase in universal distance is 1.7 per cent compared with Grid A, but of course it is nearly four times less than for the contiguous linear block.

These grids are illustrative, of course, rather than a proper test, because huge combinatorics are involved, and in complex situations the four principles will interact. For example, in Grid M, we break many lines, and also make many smaller blocks in the centre. The result is a decrease in universal distance of 1.9 per cent compared with Grid A in spite of the shortening of lines.

The law of centrality

It is now proposed that these four principles can be reduced to two formally demonstrable laws: a law of centrality and a law of compactness. The law of centrality proposes that an object placed centrally in a space will increase universal distance more than one placed peripherally. Consider again the line of n cells with an object placed somewhere along it, leaving n cells in some distribution on the two sides of the cell with at least one cell on each side.[33] Wherever we place the object, D for one side of the line must be equal to D for the other, since each cell acquires one d for each cell on the other side of the object (see Figure 17.18, mid-left). For example, if there are x cells on one side of the line and y on the other, then on one side D will be $x*y$ and on the other $y*x$. To establish D, then we need only establish it for one side of the line, since we may then multiply by 2 to get the total for the whole line. We therefore work by calculating D as the sum of ds for one side of the line.

Suppose then that the object is placed centrally on the line. It will then have equal numbers of cells on either side. Let $m = (n - 1)/2$ be the number of cells on each side of the object. Each of m cells on one side then requires one deviation to go to each of cells on the other, giving a total of $m*m$ or m^2 deviations for each side. The total deviations, D, for the line with a centrally placed object, c, is then $2m^2$ or m^2 for each side:[34]

$$D(c) = 2m^2 \quad (1)$$

Now move the object one cell sideways. The total deviations for one side will then be $(m-1)$ $(m+1)$ and for the other $(m+1)$ $(m-1)$ or $2(m-1)$ $(m+1)$ for the whole line. Now, $m^2 > (m-1)$ $(m+1)$ is a necessary inequality, as for example, $3^2 > 2*4$ or $4^2 > 3*5$. Similarly, $(m - 1)$ $(m + 1) > (m-2)$ $(m + 2)$ is a necessary inequality, as, $2*4 > 1*5$ or $3*5 > 2*6$. In general: for $D(c_1, c_2 \ldots c_n)$ representing steps away from the central location:

$$m^2 > (m-x)\ (m+x) > (m-(x+1))(m+(x+1)) \quad (2)$$

with $x=0$ for the central object case $(m-0)$ $(m+0) = m^2$. It follows that the greater the x, that is the farther the object from the centre, then the smaller is the product of $(m-x)$ $(m+x)$. In other words, the farther the object from the centre, the lower the total D.

One way to think of this is geometrically. The perimeter of a rectilinear shape is the sum of its sides. Holding the perimeter of the shape equal, the area of the shape, which is the product of its longest and shortest side, is maximised when all sides are equal and reduces as we shorten one side and lengthen the other, e.g., $4*4 = 16$, $5*3 = 15$, $6*2 = 12$ and $7*1 = 7$. In a sense, then, the law of centrality replicates the behaviour of area–perimeter ratios, even though we are dealing with linear effects.

In practical terms, this means that in a growing cellular aggregate such as a settlement, when faced with a choice of placing an object somewhere between two other objects, we should always place it close to one object and far from the other if we wish to minimise the gain in universal distance in the system as a whole. This means that gain-minimising decisions will always tend to create long and short lines rather than lines of similar length. This is clearly the case where a partitioning is made along an existing line, so that the two newly created sub-lines are colinear (that is, share the same alignment). However, it is clear that it will also be the case for lines that are not colinear. The lower depth gain from a long and short line is not created by the rule but by the situation created by the rule. It has arisen from the intrinsic properties of a longer and shorter line compared to a pair of equal lines, and the lower depth gain from the long and short line, results from the existence of the longer line, even when offset against a shorter line which was the by-product of its creation. If longer lines are beneficial even when offset by a colinear short line, it follows that a longer line will be beneficial anywhere even when it is not so offset by a colinear short line. It follows that to minimise depth gain in a system we should always conserve longer lines at the expense of shorter lines. This is the 'principle of extension' – always conserve long lines and partition shorter ones – and it thus follows as a corollary of the law of centrality. A second corollary is that placing two objects equidistant from each other and from other objects will increase universal distance more than placing them either close to each other or close to other objects, since the former will create many equal short lines, while the latter will create some longer and some shorter lines. In general, we may say that placing objects in proximity to each other increases universal distance less than placing them farther apart.[35]

The law of centrality thus addresses the fundamental spatial problem of settlement: how to aggregate built forms in such a way as to preserve the interaccessibility which is potentially interrupted by those built forms, and how to maintain this as the settlement grows. It leads to a fundamental idea in the generation of settlement: that to minimise universal distance in the system (that is, to maximise metric integration) the fundamental strategy must be always to conserve longer lines, if necessary at the expense of creating other short ones.[36]

The law of compactness

The law of compactness proposes that the more compact an object or group of objects, that is, the more its shape approximates a circle (or for practical purposes a square), then the less the increase in universal distance in the surrounding space. This may be shown by first considering the effect, as before, of placing an object on a line of n cells. We know that the maximum increase in universal distance for each side is $d*m^2$ (where $m = (n - 1)/2$) for the case where the object is placed centrally. If we then place a discrete object on a second line with at least one line between the new and old line, then the gain on each side of the second line will also be $d*m^2$, since the objects do not affect each other (see chapter 8 of *Space is the machine*, for a discussion of the case where lines are neighbours). In general, the depth gain for single discrete objects placed centrally on distinct and separate lines will be $n*2*d*m^2$, where n is the number of lines with objects. The rate of increase is therefore linear.

Now suppose that the objects are placed contiguously on neighbouring lines. This creates a more complex situation in terms of depth gain, which is illustrated in Figure 17.20. As we can see, depth gain is least at the edges and greatest in the centre. With m being the length of the line blocked and n the length of the partition (= the number of lines blocked), the depth gain can be calculated by the finite series which describes the number of deviations in one direction:

$$D = n^2m^2 + (n-2)^2\ m^2 + (n-4)^2\ m^2 + \ldots + m^2 \quad (3)$$

which gives a third-order polynomial function for the increase in universal distance with either increased partition length or line length. It can then be compared to the linear rate for discrete cells. If blocks are discrete, then universal distance increases linearly, and if contiguous, the increase is a third-order polynomial function with increasing contiguity. This demonstrates the old 'principle of contiguity'. However, as we will see below, we must also unify this with the idea of compactness.

Consider the effect of an aggregate of objects forming an overall shape placed on a regular grid of lines. The shape will increase universal distance in two directions in the grid, which we can think of as horizontal and the vertical. Holding m, the length of the line on either side to the shape, constant, the increase in universal distance in one direction will be a third-order polynomial function of n, the number of contiguous cells composing that face of the shape. Alternatively, we can hold n constant and vary m, with the same result. These calculations will not be affected by the number of cells on the adjacent side of the shape, since these will only increase universal distance in the other, orthogonal direction. The overall increase in universal distance resulting from the imposition of the shape of the grid will then be the sum of the effects on each direction of the lengths of the two different faces of the composite object blocking that direction calculated by formula (3) applied independently to both directions.

Suppose then that the sides are equal, that is the object is maximally compact, say $2x2$. Holding m constant at, say, 3, the gain in universal distance will be $2(n^2m^2) = 2(2^23^2) = 72$ for each direction (made up of the two half-lines), or $4(n^2m^2) = 144$ for the whole object. Now alter the shape of the object to a $1x4$. The gain in the vertical direction will now be $(4^23^2) + (2^23^2) = 180$ for each half-line, *2 for the pair of half-lines = 360. That in the horizontal direction will be $2(1^23^2) = 18$ for the pair of half-lines. The total gain is then 378 compared to 144 for the square object. In fact, if we reduce the object to a linear block of three cells, then we have $2((3^23^2) + (1^23^2)) = 180$ for the vertical direction and $2(1^23^2) = 18$ for the horizontal direction, giving 208, which is still greater than 144.

The reason for the increase is simple. Since m is constant, n is the only variable in Equation 3. When the block is square then $D = 2n^2$. However, if we replace the square object with a rectangular object, say, $(n-1)$ on one side and $(n+1)$ on the adjacent side, then all we need to know is the relation between $(n-1)^2 + (n+1)^2$ for the two unequal half-lines of the rectangular object and $2n^2$ for the two equal half-lines of the square object. Since $(n-1)^2 + (n+1)^2 = (n^2 - 2n + 1) + (n^2 + 2n + 1) = 2n^2 + 2$, it follows that $(n-1)^2 + (n+1)^2 > 2n^2$ and that in general:

$$(n-x)^2 + (n+x)^2 < (n-(x+y))^2 + (n+(x+y))^2 \qquad (4)$$ [37]

From this it follows that a compact form will always generate less depth gain than an elongated form of equal area, and that the difference increases rapidly with increased elongation. As with the law of centrality, a simple geometrical idea underlies the law of compactness.

Impact of the laws on the basic generative process

How then do the spatial laws impact on the basic generative process? We have already seen that the social forces driving settlement formation are dual, with a residential component, driven by socio-cultural forces, and a public space component, driven by micro-economic forces. These correspond to a duality in the settlement form itself, with the invariant deformed wheel global structure formed by the public space process and the culturally specific interstitial local background areas formed by the residential process. We also note that there is a duality in the spatial laws, in that the compactness law addresses the physical component of the settlement, that is the size and shape of aggregate objects (that is, blocks), while the centrality law addresses the spatial component, that is length of lines, distance of objects from each other and so on. We recall that the output of the basic generative process in Figure 17.12 was deficient in both respects: blocks were overly varied in their shape and lines were insufficiently varied in their length. Our task was to explain the differences between the computer-generated model and the real cases by showing how the dual social processes impacted on the basic generative process through the intermediary of the spatial laws.

Two conjectures can now be proposed. The basic generative process guaranteed interaccessibility but it did not specify its degree or type, that is, it did not specify a more or less integrated process or a particular local geometry. To control this, one would need, in the first instance, to set a parameter for the compactness law regulating the size and shape of blocks, by specifying, for example, for how long, and where, one could continue adding to an existing block, and when a new one had to be started. Such a parameter would in effect specify how the compactness law would influence the pattern and degree of universal distance in the background structure of the system in general. The first conjecture is that it is this local interaccessibility parameter controlling the generic block structure and operating through the compactness law that is set by the residential process and its socio-cultural drivers. It is through this that the characteristic local geometry of space is created in the first instance in the background residential areas of the settlement. Where this is set differently by different cultures, we find the kind of differences noted in the different parts of Nicosia (Figure 17.11). Where it is more homogeneous, we generate the kinds of generic regional differences in axial geometry that are indexed in the geometric and syntax values (Table 17.1) set out earlier.

The second conjecture is that with the growth of the settlement (and, in fact, in quite early stages) the public space process, led by micro-economic activity, sets a global interaccessibility parameter working through the centrality law. Since micro-economic activity is by nature integrative, this is not a variable, but a constant. Its effect is always to seek to conserve longer lines and to use these to minimise universal distance in the larger-scale system. Since the effects it seeks are spatial, it operates directly on space and therefore works through the centrality law. The public space process thus tends to generate the local-to-global deformed wheel structure at whatever level of the settlement it is applied, including, where it is operative, local area structures. However, this is not all the micro-economic process does. In its loci of most concentrated activity, it will generate not a linear system that minimises universal distance in the system as a whole, but a locally intensified grid that minimises movement distances from all origins to all destinations in the local region (see, for example, the central area of Konya in Figure 17.10).[38]

Looking at Konya, we can now see the settlement plan in a new light. We can see how spatial laws, driven by the dual process, have created the key features of the layouts: a deformed wheel global structure, an intensified grid forming the hub of the wheel, and the background of residential areas. However, there is an important respect in which the processes that create these patterns can be seen as a single process. The operation of the centrality law is *dual*, in that it creates both integration and segregation. In this, it is capable of reflecting, *in itself*, the fundamental duality of the socio-cultural and micro-economic processes. The socio-cultural process, which creates the larger areas of background space in the city, is always a matter of imposing some restriction on integration and the natural co-presence that follows it through movement, while the micro-economic process operates of necessity by always maximising integration (minimising universal distance) in order to maximise natural co-presence in its spaces. The micro-economic process therefore naturally occupies that part of the duality of the law of centrality which generates the longer lines and the essential structure of the settlement, while the socio-cultural process equally naturally occupies the obverse side, the production of a larger number of shorter lines which construct the less integrated background of mainly residential space in the interstices of the global structure. Through the dual nature of the centrality law, then, the dual process acquires a single expression.

These conjectures require, of course, a whole research programme to test them, involving both simulations of settlement growth and the analysis of real cases. However, some useful preliminary indications have been gained by some simple experiments on the impact of the centrality law upon the basic generative process (using at this stage a manual process). For example, once we know the law of centrality, we can use it to maximise universal distance in a restricted random process by having a rule which requires the blocking of the longest line whenever an opportunity presents itself. Figure 17.22 is a manually generated outcome from applying this rule within the basic generative process. The outcome pattern is primarily composed of short lines and (for the reasons given earlier) has very poor and non-urban-like local-to-global synergy of .147, about as low as it can get for a small system. It also lacks the kind of global structure typically found in settlements, although it does begin to show signs of an interesting, but overly peaked, log distribution of line lengths (Figure 17.23). In short, it shows little sign of the spatial invariants of settlement we are looking for. In some ways, it is the opposite. Suppose then that we use the centrality law in the opposite direction and set up a rule that forbids blocking a line once it has acquired a length of, say, five cells. This generates a pattern of many more long lines, as in Figure 17.24, which does have a good synergy score (0.820), but the lines do not construct a deformed wheel pattern with interstitial local areas, and as Figure 17.25 shows, the approximation of a log-normal distribution is quite poor.

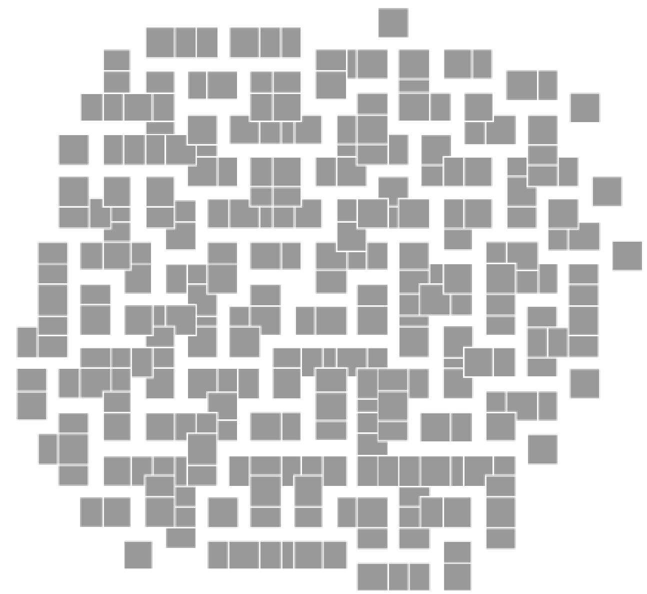


Figure 17.22 Restricted random process maximising universal distance.

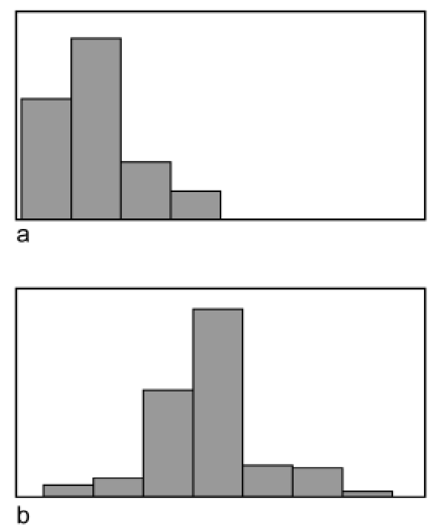


Figure 17.23 Frequency distribution of line length for Figure 17.22: (a) unlogged; (b) logged.

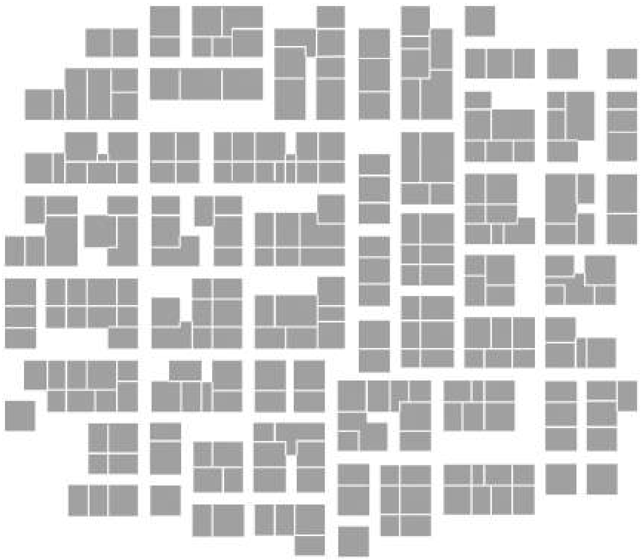


Figure 17.24 Generative process where long lines are not blocked.

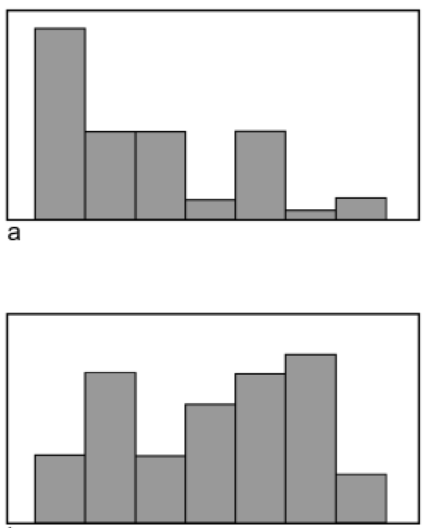


Figure 17.25 Frequency distribution of line length for Figure 17.24: (a) unlogged; (b) logged.

Suppose then we apply the centrality law in the simplest, and most localised, way by setting up a rule that says that wherever you are adding a built form to the aggregate, you have to choose a local location which preserves the locally longer line, but at the cost of continually creating shorter lines. Figure 17.26 is an outcome of such a process. Its global structure is overly biased towards the central horizontal line, but it is centre to edge, and the local areas are insufficiently structured in relation to the global core (giving it an urban synergy score of 0.729), but it does even at this stage of growth begin to look more like the log-normal distribution of line lengths, as in Figure 17.27.

This suggests that it may indeed be the duality of the centrality law in creating many shorter lines to compensate for each longer one that is in the last analysis responsible for the log-normal distribution of lines lengths in real settlements. However, although the tendency of the micro-economic process to use the longer line output of the centrality law seems to be invariant, the relations between these two aspects of the dual process should perhaps be seen as a variable. Sometimes, for example, the zones of background residential space seem to be no more than the by-product of the micro-economic process, while elsewhere – Konya would be an example – there is a conscious parametrisation of the obverse side of the dual processes to create quite substantial regions of the urban grid, sometimes quite distant from the main settlement structure. In other cases, such as London, we find the local areas are much more closely related to the global structure, more axially integrated into it, and themselves have local to global deformed wheel structures.

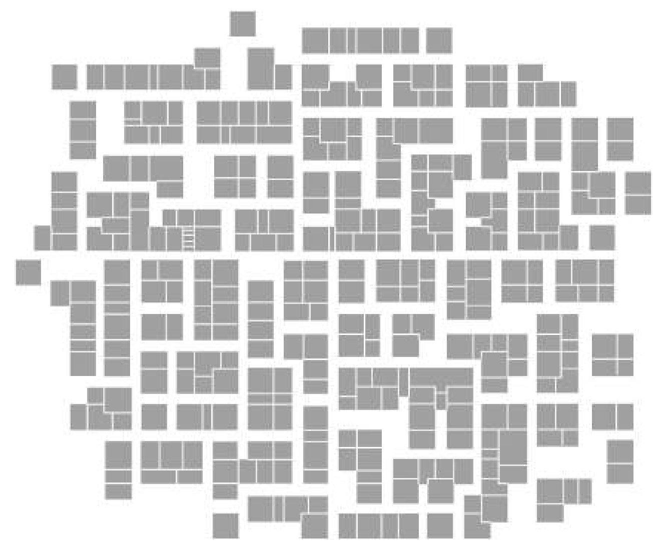


Figure 17.26 Generative process which preserves local long lines while adding short lines.

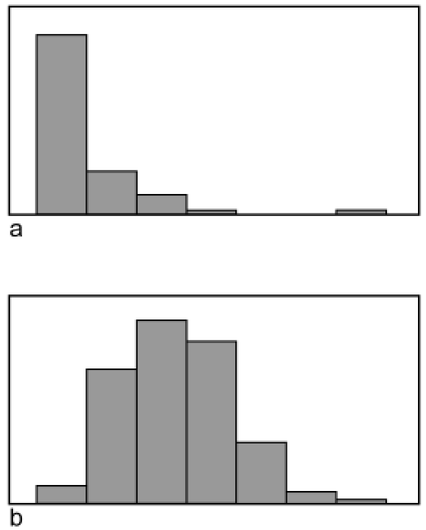


Figure 17.27 Frequency distribution of line length for Figure 17.26: (a) unlogged; (b) logged.

This kind of variation suggests a rudimentary typology of settlement forms based on the different balance between the micro-economic and socio-cultural forces. Where the economic process is dominant from the beginning, we find linear or crossroad settlements and these are usually found on major routes between larger towns, a linear town being 'global structure only'. A deformed grid town is one in which both processes run in parallel. A regular orthogonal grid town is one in which the local cultural process is in the spatial image of the global economic process, and where the whole grid is essentially a micro-economic rather than sociocultural creation, as can reasonably be said both of medival planted towns and early American grids.[39]

We may then be within striking distance of grasping aspects of the pervasive logic by which apparently different social forces generate invariants in their settlement patterns as well as the more obvious differences. The key issues are the parametrisation of the cultural process which defines the local spatial geometry, and the balance between this and the emerging micro-economic process as the settlement grows. In the early stages of growth, the local socio-cultural process guarantees interaccessibility in the emerging settlement pattern but little more. It sets a parameter which by deciding the degree or ease of interaccessibility (that is, more or less universal distance) specifies the local geometry of the settlement, covering both line length, angles of incidence and block size – all factors in interaccessibility. With growth, the universalistic and therefore globalising micro-economic process increasingly interposes on this process a simple depth-minimising mechanism for each built form placing decision: conserve long lines, if necessary at the expense of creating many shorter lines. This will have the effect of generating a pattern of a few long lines and many short lines, and because the choices are regional this will be the case at every level – that is, this process will generate the pervasive log-normal distribution with a few long lines and a large number of short ones at every level. Changes in this fundamental pattern of growth will reflect essentially the changing balance between micro-economic and cultural forces, and this may (as historically in London) alter with the passage of time, with each alteration leaving its mark on the settlement geometry.

However, the core issue is that the inherent duality of the spatial law of centrality is able to reflect the duality of these potentially conflicting social forces and turn what is initiated as a dual process into a single process by which the locally highly differentiated and globally highly structured pattern of urban space come into being.

A reflection

The deformed wheel structure with its interstitial areas – the classic, although not the only, urban form – seems thus to be a product of an essentially metric process, optimising metric integration in some aspects, restraining it in others. Some may have noticed that this leads to a difficult question. Why should we continue to regard axial maps as topological structures, to be analysed through their graphs, when we have shown that they are generated through an essentially metric process? Would we not be likely to arrive at a better picture of the city if we subjected the axial map to metric analysis? It has already been suggested that the intensified grids found in centres and subcentres are best understood through metric integration analysis.[40] Is it not time to subject the axial map as a whole to such an analysis, or at least to a metrically sensitive analysis? In this way, we could surely counter one of the main objections to the axial map as a basis for graph analysis: that the nodes of the graph represent unequal elements.

The problem is that as soon as we introduce a metric dimension to an axial map – whether by using an analysis based on metrically uniform elements, or by weighting, say, line segments for length – configurational analysis produces not an enhanced version of the kind of picture given by the line graph analysis, but a very different picture: one that essentially picks out geometric centrality in the system, as we saw when we used more or less uniform line elements in the pseudo-system shown in Figure 17.15. If we applied this to a city like London, it would have the effect that a short alley off Oxford Street would seem to be more integrated than, say, the Holloway Road. In one sense it is, of course, since it is closer to the geometric centre of London. However, in a more important sense we would seem to be losing one of the most important aspects of the integration analysis of the urban system: the substitution of a picture of geometric centrality by a picture of centrality in the line topology, one that identifies geometric centrality but then draws it out towards the edges of the system in all key directions, and even including parts of the edges of the system.

The question is then: which is the true picture? Is the one brought to light by the radius *n* analysis of the line graph in some sense identifying properties that are truly of the nature of the urban system and essential to its functioning? One thing is clear. Metric analysis of a large-scale system is very much poorer in its capacity to post-dict the movement structure. In experiments carried out in 1986 on axial maps whose segments were weighted for length and used as the units of analysis, this very propensity to assign too high a movement prediction to lines adjacent to strong lines and too low a prediction to syntactically stronger but more remote lines destroyed the normal approximate agreement between integration and movement.[41] This suggests that the axial map, analysed as a line graph, might after all be capturing something that is of the essential nature of the urban system.

What can this be? There are two aspects to a possible answer: one substantive and to do with urban reality, the other cognitive and to do with how we interact with urban reality. Substantively, the empirical effect of the line inequalities in the urban system is to create a disjunction between geometric centrality in the system and topological centrality in the line map. In effect, centrality is topologically stretched from the geometrical centre to form links with the edge in all directions. In doing so it also structures the object by creating a relation between the local and the global organisation. The benefits of these are obvious enough: strangers are provided with easy-to-read routes from edge to centre and out again, and the system acquires local-to-global intelligibility and synergy. In contrast, it is easy to see that a system without the line inequalities in the right place, and of the right type, will degenerate into a labyrinth.

In fact, in terms of the micro-economic processes that create the deformed wheel structure, we find an even stronger argument when we consider the settlement, not in isolation, but as part of the wider system of settlements. Figure 17.28 illustrates the 'paradox of centrality'.[42] On the top left are four notional settlements, each with its own internal integration core. But when (on the top right) we join them into a single system and analyse the integration pattern for the system of settlements, we see that integration shifts to the edges of the settlements. Bottom left, if we link centre to edge lines to each other, the core penetrates into, but not across, the settlements. Bottom right, if we extend the centre to edge lines into neighbouring settlements, the internal cores of each are nearly restored. Clearly, if we consider each settlement on its own, then the internal pattern of integration will approximate the internal movement structure, while if we consider them as a system of settlements, the edge pattern will reflect movement in the overall system.

This is of course exactly what happens in real settlements. Movement patterns invariably have a local aspect and a global aspect, the former reflecting circulation within the system, the latter movement in and out of the system. Insofar as movement is driven by the micro-economic process, it generates both

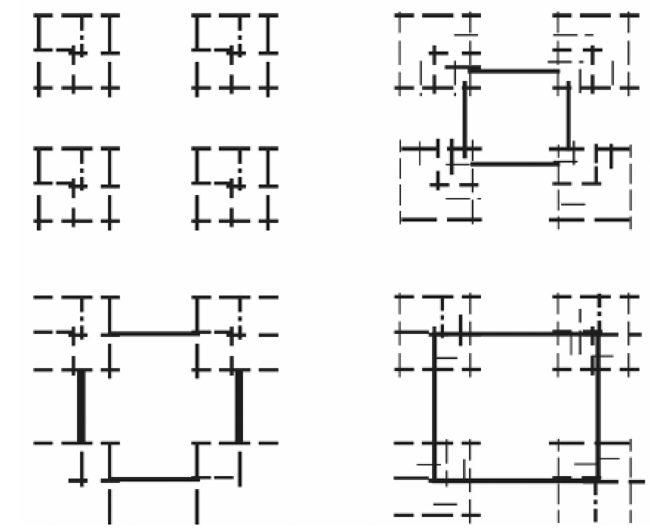


Figure 17.28 Four notional settlements and ways of connecting them, with resulting integration cores.

the intensified local grids of the centres and subcentres by reflecting the need to minimise distance from all points to all other points within the zone, and the linear links from the local to the global scale of the settlement, reflecting the need to minimise distance from certain points to certain others at the larger scale,[43] including into and out of the system. Over time, this tension between the internal and external movement economies of the settlement is the fundamental reason why centrality tends to shift towards the edges of the settlement, unless strenuous efforts are made to inhibit it.

The deformed wheel structure is the key mechanism for this inhibition. In Figure 17.28, we draw four hypothetical settlements, each with a 'deformed wheel' core linking edge to centre (top left). Top right, we connect the edge lines of the four. The overall core goes to the edge of each settlement. Bottom left, we link centre to edge lines to each other. The core penetrates into, but not across, the settlements. Bottom right, we extend the centre to edge lines into neighbouring settlements. The internal cores of each are nearly restored.

Cities as discrete geometries

The second reason why we might suspect that the axial map captures essential properties of the urban system is cognitive. The analysed axial map seems to approximate the intuitive picture we have of an urban system to an unexpected degree. A simple reason for this would be that human beings are excellent judges of simple linear distances when, for example, throwing a stone or a spear, or a ball of paper into a waste paper basket. However, this comparatively secure judgment of distance quickly breaks down when the system becomes nonlinear and involves changes of direction. This would make simple sense in evolutionary terms. Distance is a comparatively sophisticated and recent concept, and there is no obvious reason why we would expect to judge it as well in the highly nonlinear situations created by human settlement as when we are dealing with distance as a simple extension of bodily reach.

Complex spatial systems seem then to be dealt with cognitively through something more elementary. What might this be? The obvious candidate is discrete geometry: that we cognise complex spatial systems like cities as assemblages of interrelated geometrical elements rather than as complex patterns of metric distance.[44] Discrete geometry is the application of the techniques of discrete mathematics such as graph theory to systems of discrete geometric elements, such as lines, convex spaces and visual fields. Space syntax, we now can say with hindsight (there was not much discrete geometry about when we started), is the application of discrete geometry to architectural and urban systems, considering these first and foremost as systems of space.

If our cognitive representations of complex space are indeed discrete geometrical, then the strongest candidate as the element in the discrete geometry would be the line. Lines have the two key properties of being both very simple and very global. All we need to know is how far we can see from a point. Put more theoretically in terms of the city as a total visibility field, we can follow Penn (and adapting Peponis's beautiful concept of informational stability as those regions in a spatial system that do not change topologically with movement[45]) in arguing that a linear clique (a set of points which can all see each other) preserves informational stability for longest for moving individuals and thus offers the most economical – although not the most complete – picture of an overall system.[46] Other discrete geometrical representations, such as visibility graphs[47] give a much more complete account of the complexities of urban space, but it is not obvious that they would form the basis of a cognitive representation of the city as a whole. There is too much local information for the global picture to be clear. An axial map maximises local simplicity as a means to picturing global complexity. With visibility graphs, it is the other way round. Analysis of how we give directions in complex spatial systems suggests that the axial maps may not be too far from the way we represent them to ourselves, that is as a matrix of lines where changes of linear direction are the key items of information that become organised into the whole picture.[48]

If we intuit the spatial structure of the city as a *discrete geometry*, then it is reasonable that we should analyse it by treating the discrete elements as the nodes of a graph. We are tempted to add to this that we represent the urban system to ourselves not simply as a discrete geometry, but as a simplified discrete geometry, in the sense that a series of near straight lines of the kind that are commonly found in cities are internally represented as a line, so that the whole system comes to resemble an approximate grid.[49] If this is the case, then it would be no more than a case of the imposition of a Euclidean framework on non-Euclidean inputs' as argued by Peterson and her colleagues:[50]

> . . . the hippocampus appears to impose a Euclidean framework on non-Euclidean inputs O'Keefe and Nadel (1978), who see in this process an instantiation of a Kantian *a priori* notion of absolute space we propose that in 'distorting' the sensory inputs, these spatial maps impose an order and a structure that our spatial conceptual representations require.[51]

Since it is also the line topologies that seem to correlate with movement in the different parts of the system, it seems hard to avoid the conclusion that the line representation of the city is not just a convenient simplification but something that touches the essential nature of the city. This does not mean that it cannot be improved or broken down more than it is now. However, it does seem likely that any future configurational analysis of the large-scale structure of cities will need to include some representation of its linear dimension as currently expressed, although perhaps crudely, in the axial map.

With or without the axial map, this account of how urban space is generated has unavoidable implications for how we model the city. Models in the past have used the fundamental concept of mass and the Newtonian mathematics of gravitational attraction as the guiding theoretical entities. The integration equations play the same role in configurational models as the Newton equations do in attraction-based modelling. But they do so on the basis of a discrete geometrical representation of the spatial structure itself, one that seems to engage the key role of visibility in how we cognise and interpret our surroundings. This has a further implication: as Chiron Mottram has argued, configurational models are light-based rather than mass-based: they reflect the world we see rather than the world of distance and mass.[52] The question is: how far do these cognitive realities intervene in the functioning of the urban system? On the evidence so far, it seems unlikely that we can arrive at a theoretical model of the city without them.

Notes

1. B. Hillier, 'The architecture of the urban object', *Ekistics* 56 (1989): 5–21; B. Hillier, A. Penn, J. Hanson, T. Grajewski and J. Xu, 'Natural movement: Or, configuration and attraction in urban pedestrian movement', *Environment and Planning B: Planning and Design* 20 (1993): 29–66; B. Hillier, 'Centrality as a process: Accounting for attraction inequalities in deformed grids', *Urban Design International* 4 (1999): 107–127.
2. M. Batty, 'Foreword', in *The mathematics of urban morphology*, edited by L. D'acci (Cham, Switzerland: Birkhäuser, 2019), vi–x.
3. A. V. Moudon and C. Lee, 'Urbanism by numbers: A quantitative approach to urban form', in *Making the metropolitan landscape: Standing firm on the middle ground*, edited by J. Tatom and J. Stauber (New York: Routledge, 2009), 57–77, p. 73.

4. See definition of 'deformed wheel' in Hillier, 'The architecture of the urban object', p. 10
5. Martin's 'The grid as generator', republished in *ARQ* in 2000 for its importance in urban studies, was a chapter in L. Martin and L. March (eds), *Urban space and structures* (Cambridge: Cambridge University Press, 1972), 6–27.
6. After Martin, 'The grid as generator'.
7. Hillier et al., 'Natural movement'; B. Hillier, 'Space as a paradigm for describing emergent structure in strongly relational systems', in *Proceedings of the Second International Space Syntax Symposium*, Brasília, edited by F. de Holanda, L. Amorim and F. Dufaux (1999), pp. 1–16; Hillier, 'Centrality as a process'.
8. An 'axial map' is the least set of longest lines of direct movement that pass through all the public space of a settlement and make all connections.
9. Hillier et al., 'Natural movement'.
10. The 1993 paper dealt only with 'global' or radius-n analysis, but a series of studies since then have shown that 'local' or radius-3 integration is normally a better predictor of pedestrian movement.
11. B. Hillier, 'Cities as movement economies', *Urban Design International* 1 (1996): 49–60.
12. B. Hillier, 'Centrality as a process'.
13. As reviewed in J. V. Skaburskis, 'Territoriality and its relevance to neighbourhood design: A review', *Journal of Architectural Research* 3(1) (1974): 39–44.
14. B. Hillier, *Space is the machine: A configurational theory of architecture* (Cambridge: Cambridge University Press, 1996).
15. In this sense, the argument is still within the spirit of the theoretical framework set out by Martin and March in *Urban space and structures* in 1972.
16. The question 'What about planned towns?' may of course be raised here. However, in the great majority of cases the planned element is only the first stage of an urban growth process that then will be subject to the same lawful influences as cities which have grown through a distributed process.
17. In some cases, such as Chicago and Amsterdam, we find a loglognormal distribution. However, the difference between a log and a loglog distribution is much greater than that between an unlogged and logged distribution, so these differences are not pursued here.
18. F. de Holanda, 'Exceptional space' (PhD diss., University of London, 1997). Editors' note: Holanda's dissertation has since been published as *Exceptional space* (Brasília: FRBH, 2011).
19. Hillier, 'The architecture of the urban object'.
20. Hillier, 'Centrality as a process'.
21. This phenomenon is also found in subcentres. In 'Centrality as a process', it was argued that wherever movement is convex and circulatory (that is, moves around in a locally two-dimensional grid as, for example, in a shopping centre) rather than linear and oriented (as in moving through an urban grid from an origin to a destination), then metric integration was the key property in understanding both the movement pattern and the type of spatial configuration that tended to emerge under these conditions.
22. A. Loumi, 'Spatial configuration of towns in North Africa' (PhD diss., University of London, 1988); K. Karimi, 'Continuity and change in old cities: An analytical investigation of the spatial structure in Iranian and English historic cities before and after modernisation' (PhD diss., University of London, 1998); Hillier, 'Centrality as a process'.
23. B. Hillier, and J. Hanson, *The social logic of space* (Cambridge: Cambridge University Press, 1984).
24. In the version of the process set out in *The social logic of space*, the open space of the dyad was the same size as the built cell. In the version shown here, this has been retained, but the built cells have then been expanded without expanding the open spaces, with the effect that the scaling of open spaces and buildings approximates real systems more closely.
25. Holanda, 'Exceptional space'.
26. Old Paranoá has now been pulled down by the planning authorities and replaced by a much more regular settlement.
27. The 'intelligibility' correlation between [axial] connectivity and global integration is used here rather than the 'synergy' correlation between local and global because the systems are initially too small to respond realistically to local integration analysis. The argument would, however, hold up for 'synergy' analysis.
28. For an account of the idea of 'universal distance', see *Space is the machine*, chapter 3. Universal distance is probably the most fundamental concept in space syntax. It can be applied either metrically or topologically, and allows the redefinition of an element in a system as no more than a position from which the rest of the system can be seen, thus 'nearly dissolving the elements'. (See Hillier, 'Space as a paradigm'.)
29. Set out in chapter 8 of Hillier, *Space is the machine*.
30. Hillier, *Space is the machine*, chapter 8.
31. On reflection, what were noted in *Space in the machine* were empirical regularities, since no theoretical account was offered as to why they should be so.
32. Hillier, 'Centrality as a process'.
33. The argument is best understood by supposing that n to be an odd number [eds.].
34. Editors' note: It is important to notice the difference between the total number of deviations, *D*, and the total depth gain. To get the total depth gain we need to multiply the number of deviations by the magnitude of each deviation. In the examples of Figure 18, d=2. Thus, to arrive at the total depth gain recorded in Figure 17.18 we multiply the total number of deviations by 2.
35. This may be demonstrated with greater clarity in an unbounded system such as a torus. Consider two objects placed on the surface of a torus. The two objects have two distances from each other: a distance from one face of the object to the nearest face of the other object; and a second distance from the opposite face of the first object to the opposite face of the second object around the other side of the torus. These distances may either be equal or different. If they are equal – that is the objects are the farthest possible distance apart in any direction – then the law of centrality shows that the increase in universal distance is maximal. The more the objects are moved together to create a nearer and a farther distance, the more we have a shorter and a longer distance and therefore the less the increase of universal distance. It follows that placing objects close to each other in an unbounded system increases universal distance less than placing them farther apart.
36. Also, variation in the scale of the attached space will have substantial effects on the axial map.
37. Editors' note: x and y are not clearly defined here. x seems to stand for the increment by which one side of the composite and initially square object is elongated and the other shortened. y seems to stand for an additional increment. The equation simply states that a more elongated object causes greater depth gain than a less elongated one.
38. B. Hillier, 'The hidden geometry of deformed grids: Or, why space syntax works, when it looks as though it shouldn't', *Environment and Planning B: Planning and Design* 26 (1999): 169–191; Hillier, 'Space as a paradigm'.
39. Such 'economic' grids need, however, to be distinguished from the grids of administrative, garrison or ceremonial towns which characteristically are not pure grids but interrupted grids in which many lines, including some major lines, are interrupted by the facades of major public buildings at right angles.
40. Hillier, 'Centrality as a process'.
41. B. Hillier, 'Appendix: The distance factor', *Spatial configuration and use density at the urban level: Towards a predictive model*, final report to the Science and Engineering Research Council (London: Unit for Architectural Studies, Bartlett School of Architecture and Planning, University College London, 1986).
42. The term was first coined in chapter 9 of Hillier, *Space is the machine*.
43. Hillier, 'The hidden geometry of deformed grids' and Hillier, 'Space as a paradigm'.
44. J. Goodman and J. O'Rourke, *Handbook of Discrete and Computational Geometry* (New York: CRC Press, 1997).
45. J. Peponis, J. Wineman, M. Rashid, S. H. Kim and S. Bafna, 'On the description of shape and spatial configuration inside buildings: Convex partitions and their local properties', *Environment and Planning B: Planning and Design* 24 (1997): 761–781.
46. A. Penn, 'The role of spatial configuration in spatial cognition: Or, why the axial line?', in *Third International Space Syntax Symposium*, edited by J. Peponis, J. Wineman and S. Bafna (Atlanta, GA: Georgia Institute of Technology, 2001), pp. 11.1–11.17.
47. A. Turner, M. Doxa, D. O'Sullivan and A. Penn, 'From isovists to visibility graphs: A methodology for the analysis of architectural space', *Environment and Planning B: Planning and Design* 28 (2001): 103–121.
48. Hillier, 'The hidden geometry of deformed grids' and Hillier, 'Space as a paradigm'.
49. Hillier, 'The hidden geometry of deformed grids' and Hillier, 'Space as a paradigm'.
50. Commenting on J. O'Keeffe and L. Nadel, *The hippocampus as a cognitive map* (Oxford: Clarendon Press, 1978).
51. P. Bloom, M. A. Peterson, L. Nadel and M. F. Garrett, 'Space and language', in *Language and Space*, edited by M. A. Peterson, L. Nadel, P. Bloom and M. F. Garrett (Cambridge, MA: MIT Press: 1996), pp. 553–577.
52. Chiron Mottram, 2001, personal communication.

18 Society seen through the prism of space (2002)

Hidden connections across scales: Anthropologically charged configurational thinking as a cure for localism. Introduction to 'Society seen through the prism of space'

Vinicius Netto

'Society seen through the prism of space' was published as a journal article in 2002. As its subtitle ('Outline of a theory of society and space') clearly states that it attempts to understand what holds a social world together. Back then, Lyotard's post Modern critique of encompassing theories had already spread like wildfire across the social sciences, and approaches focused on large-scale processes were labelled as 'grand narrative' and declared superseded.[1] Who would need another take on broad social and material forces? It turns out we still needed them. Many felt that, as a technologically changing global urban society, we were moving into uncharted territory – or no territory at all! As information technologies like the Internet and mobile communication networks were reaching ubiquity, the very material foundations of social life and practices that led to the creation and growth of cities throughout history seemed shaken. Fully interacting from far away, humans could finally break up the centripetal forces that led to urbanisation and live scattered in post-urban formations. Or so the then fashionable fears of 'the death of distance' had us believe. However, those apocalyptic tales of the end of cities were not mirrored in our daily urban experiences. Accordingly, on a closer empirical look, cities were not breaking up to give way to some new spatiality. Cities were still growing strong, as rigorous works like Saskia Sassen's were showing.[2]

So that was the context. As theorists – one at the top of his intellectual abilities, another a newcomer, a theorist-in-the-making – our initial aim was to understand why cities resisted their announced demise. Jumping 20 years into the present, one can still feel compelled to know why cities have withstood even a new global pandemic that has finally led many of us to digitally cooperate across distances. Back at the turn of the millennium, I was puzzled by the nature of internal structural changes within cities as 'systems of systems' with distinct materialities and temporalities. I asked whether the different systems that make up cities might not simply diverge; could, for example, a fast-changing social system push slow-changing urban spaces into obsolescence? Hillier argued that the kinds of problems that he and I confronted seemed to be, at their heart, the same problem: how society has space at its core for its own reproduction; and, taking a further step in configurational reasoning, how space is overcome by the creation of social connections beyond local events, across scales – through changing forms of relationship between technology, practices and space, all intertwined in their historical transformations. Such a problem required a 'space-first' approach structured along a series of questions: some addressed the nature of society: 'what kind of system could a society be?'; 'why should human societies be shallow graphs?'; 'what is an urban society in the space of possible societies?' Other questions addressed society as a material system: 'are there hidden configurational variables?'; 'does society have necessary spatial dimensions?' Finally, there were questions about the future of cities: 'will technological and social changes lead to a radical change in the demands that society places on space?'; 'are we witnessing a radical transformation of cities or even the end of cities as we have known them?' Hillier's curious mind knew how to formulate deep-rooted questions and how to place them within a single and coherent conceptual framework.

Of course, his thinking was supported by research findings on the relationship between the morphology of space and the distribution of movement and co-presence as critical features in social reproduction. At the same time, it called for progress in new directions. If indeed space creates opportunities for serendipitous encounters and, through them, for new interactions, collaborations or exchanges, how do we bridge the gap between the description of morphologies of movement and co-presence and the documentation of unfolding networks of productive social interactions? These new questions indeed preoccupied me in subsequent work.

Returning to the time when this paper was written, one of my favourite insights triggered by this anthropologically charged radical network thinking leads to a surprising isomorphism between a society and its members. As we justify a social connection graph taking different individuals as 'the root' or 'carrier', how do we conceptualise the interdependence between individual social identities and larger social formations? What an individual is can be understood as the graph of society seen from a particular point of view – a counter-intuitive isomorphism with beautiful ethical implications.

Propositions of this kind might never be fully empirically demonstrable. Still, they have an irreplaceable role: to explore the power of theory to make us think and imagine – and amplify our view of reality. That is what this paper invites us to do: to think systemically but also imaginatively about how we manage to create immensely complex social and material worlds.

So, do we need general theories like the one sketched in this paper? We still do. We will need them as long as we wish to understand how local social processes intertwine and how large-scale processes emerge to connect localised events and situations through space and time. And we will need them as long as we live in a changing world.

Acknowledgement

I thank Professor John Peponis for our stimulating discussions and for his insightful comments.

Society seen through the prism of space

Outline of a theory of society and space

Bill Hillier and Vinicius Netto

Originally published: Hillier, B. and Netto, V. 2002. 'Society seen through the prism of space: Outline of a theory of society and space'. *URBAN DESIGN International*, 7, 181–203.

Two questions challenge the student of space and society above all others: will new technologies change the spatial basis of society? And if so, will this have an impact on society itself? For the urbanist, these two questions crystallise into one: what will the future of cities have to do with their past? Too often these questions are dealt with as though they were only matters of technology. However, they are much more than that. They are deep and difficult questions about the interdependence of technology, space and society that we do not yet have the theoretical apparatus to answer. We know that previous 'revolutions' in technology such as agriculture, urbanism and industrialisation associated radical changes in space with no less radical changes in social institutions. However, we do not know how far these linkages were contingent or necessary. We do not, in short, have a theory of society and space adequate to account for where we are now, and therefore we have no reasonable theoretical base for speculating about the future. In this paper, we suggest that a major reason for this theoretical deficit is that most previous attempts to build a theory of society and space have looked at society and tried to find space in its output. The result has been that the constructive role of space, in creating and sustaining society, has not been brought to the fore, or if it has, only in a way that is too general to permit the detailed specification of mechanisms. In this paper, we try to reverse the normal order of things by looking first at space and trying to discern society through space: by looking at society through the prism of space. Through this, we try to define key mechanisms linking space to society and then use these to suggest how the questions about the future of cities and societies might be better defined.

> The modern city is losing its external and formal structure. Internally it is in a state of decay while the new community represented by the

nation everywhere grows at its expense. The age of the city seems to be at an end.[3]

At the turning point between the twentieth century and the twenty-first, a new kind of economy is coming into being, and a new kind of society, and a new kind of city: some might say no city at all, the end of the city as we have known it, but they will doubtless be proved wrong.[4]

Introduction and review

In our first paper to the symposium, it was proposed that the social construction of space in human settlements was mediated by spatial laws.[5] The laws were of two kinds: those by which different ways of placing buildings gave rise to different spatial configurations; and those through which different spatial configurations created different patterns of co-presence among people through their effect on movement. What were loosely referred to as different social forces then expressed themselves in space through the different requirements that each placed on co-presence. For example, residence tends to restrain and structure co-presence, and therefore to arrange buildings to achieve relatively localised and restrictive spatial configurations, while micro-economic activity tends to maximise co-presence and thus to arrange buildings to integrate space locally and globally.

This shapes a city into two broadly different spatial elements: on the one hand, a residential background of spatial areas whose spatial patterning varies with culture, depending on the way in which that culture seeks to restrain and structure co-presence between, say, inhabitants and strangers or men and women; and on the other, an interlinking global system of public space, usually expressed in the axial map as some variant on the 'deformed wheel', generated mainly by spatially invariant (in the sense of always seeking to maximise co-presence) micro-economic processes. Thus, in settlements, space operates in at least two distinct modes: one, 'conservative' and the other, 'generative'. The conservative mode restricts co-presence in order to conserve or reproduce cultural patterns; and the other generates the maximum co-presence in order to optimise the material conditions for everyday life.[6] Through this theory, we were able to suggest why large settlements, in spite of their manifest differences, tend to have certain generic similarities. They are a consequence of spatial laws mediating the relation between configurations of social activity and configurations of space.

Some theoretical implications

There is of course nothing new in either of the two socio-spatial propositions on which this argument depends: that economic processes tend to operate uniformly and culture idiosyncratically. All we have done is to suggest how the same laws give spatial expression to both and, through this, generate the basic features of the spatial layouts of settlements. However, this does raise interesting theoretical questions about the current debate on the nature of cities and their possible future, or lack of a future. First, it implies that the relation of social activities to space is *generic*, rather than *specific*. It is not this or that pattern of activity that gives rise to the durable spatial patterns that we find in cities, but the demands that different *kinds* of activity make on co-presence, which articulate the spatial laws to make one kind of space rather than another. In fact, because two sets of laws intervene between social activity and space – laws governing the emergence of spatial patterns from accumulated local actions, and laws governing the impact of those spatial patterns on co-presence – it means that the relation of society and space is *two-way generic*: generic aspects of social action relate to generic patterns of space.

This is why in general – and with important exceptions – during the life of a city, space changes only slowly while activity changes rapidly. We do not find that new phenotypical patterns of activity per se generate new patterns of space, but that new patterns of activity have a certain distribution of demands on co-presence and that, to the degree that the new distribution approximates the old, the new pattern will be absorbed into the existing urban framework with comparatively little change. The appendix to this paper outlines a case study of the City of London, drawn largely from the work of Julienne Hanson,[7] showing how radical this adaptation can be. When assessing the impact of new activities on space, then, what we need to compare is not so much the contents of new activities but the range of demands they are likely to make on co-presence. The question we must ask about the future is then: have we reached a radical discontinuity in this process of slow and fast change? Are technological and social changes now generating patterns of activity that will be incompatible with the distribution of spaces that we currently have through current patterns of urbanisation. And will this lead to a radical change in the demands that society places on space, perhaps leading to a radical transformation of cities or even, as some have suggested, to the end of cities as we have known them.

A second implication of the theory is that social forces have inherent spatiality within them, so strong and systematic that it is capable of being articulated by spatial laws, and so clear that it can be detected by the careful examination of real patterns of space. What is particularly interesting from the point of view of a theory of society and space is that the spatialities we have seen operating in cities cannot just be in the nature of things, since the city is only one spatial state of society among many others. A further question then arises: might there be other social forces with other spatialities, for example, those that give rise to non-urban patterns? And how might these be relevant to the possibility of a post-urban society?

The third implication is that space plays a *constructive* as well as a *receptive* role in shaping the forms of social action that we see in cities. The question is, is it also constructive of the underlying generic patterns of urban societies, of the genotypes of urban society, we might say, as opposed to the phenotypes? This is a legitimate question because wherever human activity has generated cities, for whatever reason, it seems also to have changed a great deal else: social institutions, lifestyles, habits of thought and even the nature of human social and individual identity. A city is both a transformation of space and a transformation of society. We do not really have a coherent theory for this, in spite of the number of social as well as urban theorists who have been concerned with it.[8] It is not logically plausible that all of these changed and that cities were built as a consequence. In the rise of cities, space and society seem, at the very least, to have changed together.

The question we are now facing then is: if space is now changing, will society also change? If we are entering post-urban space, then what does this imply for post-urban society? It is clear that we cannot hope to answer such questions simply by studying cities. We need to know what it is about societies that interacts with space and underlies the changes that historically seem to have occurred in one when they occurred in the other. We need to understand what an urban society is in the space of possible societies.

Aim of the paper

The aim of this paper is to sketch a way of approaching these questions by looking initially at space, and trying to detect society through space, in contrast with most commentaries on society and space, which typically look at society and try to detect its output in space. Here we will take a frankly spatial point of view of the same question: to look at society through the prism of space, and trying to outline a theory of society and space *seen from the point of view of space*. The text will, as a consequence, be rather thin on discussions about society in the usual sense, because the aim will be to isolate what it is about society that turns itself into space, and what it is about space that turns itself into society. Having seen the signs of inherent spatiality in social forces, we are now looking for it in society itself.

The theory of society and space sketched here – and we must emphasise that it is only a sketch – implies two main critiques of much existing theorisation. The first is that because most attempts to build a theory of society and space look at society first, the *constructive* role of space in creating the generic forms of society has not been brought to the fore, or if it has, only in a way that is too general to permit the detailed specification of mechanisms. The second is that much explicit theorising about space has succumbed in one way or another to what might, with hindsight, be called the 'myth of historical spatiality' – the idea that in the past we were much more spatial and localised than we are now, and therefore find the present strange and alienating. This myth has afflicted the spatial disciplines, academic and applied, throughout the twentieth century, and obscured the implications of a growing body of research results that have accumulated over the past few decades in fields as far apart as the study of hunter-gatherer societies, tribal societies, social networks and organisational dynamics, and many others, which have in common that they suggest that the fundamental mechanisms that operate in society are not only those that solidify local groups, but also those that create non-local networks, including those which favour the non-local at the expense of the local.[9]

These results raise – or ought to raise – a fundamental question about how we see society, in general, in a spatial context. What would be the implications for a theory of society and space of the proposition that the core mechanisms in human societies historically were not only local and spatial, but also non-local and virtual? This paper will explore this question and will suggest that if we continue to contrast our present situation with the historic past on the grounds that they were local and we are global, or that they were spatial and we are virtual, then we cannot understand what is happening how. All the evidence is that human societies were always – at their appropriate scale of course – global as well as local entities, and virtual as well as spatial entities in some 'face-to-face' sense. In fact, it is only by studying the mechanisms by which societies become virtual and non-local that we can be guided towards a theory of society and space, because the very fact that a society exists means that the interstitial spaces distancing discrete individuals and co-habiting groups from each other have been overcome. It is *how* space is overcome that is the essential linkage between society and space, and, because society can only be created by the overcoming of space, it also provides clues to the morphological dynamics of a society. It is for this reason that we find that the key items of 'social software' – that is, the rules, beliefs, values and practices that guide our space-time 'situated practices' – are those that on the one hand lead to the overcoming of distance to create relations across space, and on the other hand control the effects of lack of distance among locally proximate groups. These are the raw materials of a theory of society and space.

Difficulties of the project

How then should we seek to construct a theory of society and space? We must begin by acknowledging certain, basic difficulties in the project. Whichever way we look at it, the very possibility of a theory of society and space presupposes that the relation between the two is in some sense systematic. If it were not, then there would be nothing for theory to latch on to. However, the very idea that this might be the case raises severe difficulties. Logically, there cannot be a systematic relation between society and space unless two conditions are fulfilled. The first is that space must have, or at least be capable of having, social potentials of some kind since if it does not, it cannot embody whatever it is that society sends to it. The second is that society should have, or be capable of having, spatial necessity of some kind, since if it does not, then it cannot impose itself on space in a way in which space can receive it. For example, if society is an entirely immaterial entity – say, a consensus among individuals – then it cannot matter how it is deployed in space, since all deployments will be equal and leave the social consensus as it is.

The first of these two problems has been the preoccupation of space syntax research: to show that space, through its very form and configuration, can express social potentials, carry social contents, and through this take part in the production and reproduction of society. But how can the concept of a society contain spatial necessity? The idea seems paradoxical. If we examine less inchoate and less general social concepts such as a family or an organisation, we find that each is a structure of roles and relations, which can be drawn up in a diagram that will be the same, however it is realised in space. The space-time realisation does not affect the essential description of what the social entity is. Spatial form may affect the dynamics of a family or organisation, but it will not change its basic defining diagram. How then can society differ from these lesser social entities in having spatial necessity as part of the definition of what it is?

Let us look a little more closely at what we mean by society. It turns out that there is a spatial problem at the very heart of the concept of society, which must be solved by any social theory that includes a definition of what a society is. It is obvious enough. The individuals who make up a society are clearly well-defined space-time 'things' in the sense of being bounded and occupying a well-defined and continuous region of space-time. However, it is not clear in what sense any higher-level pattern of relations among these individuals is, or even can be, in any comparable sense an acceptable 'thing'. It lacks the very combination of space-time boundedness and continuity that allows us to identify it as a thing, in anything like the normal sense. Of course, a society is likely to occupy a territory of some kind, but this does not solve the problem. To occupy a continuous territory is not the same as to constitute a continuous space-time entity. A society seems to be composed only of freely mobile discrete individuals. If it exists at all, then whatever it is, the large-scale entity, society, is not a space-time 'thing' in any familiar sense.[10]

This is the core problem of social theory and it is a problem of space. For society to exist, the spaces between individuals and between spatially proximate groups must somehow be filled up or overcome, and a superordinate system of some kind imposed. But what kind of a system can that possibly be? We have no conception of such a system. Social theories can be seen as a range of solutions to this problem. At one extreme, methodological individualism asserts that no such superordinate entity exists, and that society can be reduced to its individuals.[11] At the other, organicism – the idea that society is some kind of organism – tries to redefine what is manifestly a spatially disconnected system as one that is, after all, spatially connected.[12]

From abstract relations to empirical configurations

What kind of a system could a society then possibly be? One line of enquiry begins by acknowledging that we are up against the ancient and deep philosophical problem: that of defining what *relations* are, as opposed to the things that are related. Are relations real in the sense that space-time 'things' are real? This is most succinctly expressed by Russell. The relation that Edinburgh is to the north of London does not seem to be a material thing, in the same sense that Edinburgh and London are material things, so we are tempted to see it as a mental construct rather than as a property of the real world. Yet within the scheme of things defined by this universe, Edinburgh 'really does' seem to be to the north of London and the relation to exist 'out there', written into things themselves.[13] So, we are tempted to assign relations to a world that is neither mental nor physical, but accessible to us through our intelligence rather than our senses. However, this means that if society is essentially a relational scheme linking individuals, and relations do not belong in this world, the consequence is that 'society' is taken out of the world of space and time and placed in another one, free from space-time, from which it can surely exercise no direct influence on this one. If there can be no theory of society and space, unless we can show spatial necessity in society, then it seems that we cannot solve the problem of society and space unless we first solve the problem of the space-time status of relations.

Now from the point of view of space syntax, this is an interesting formulation of the question, because whatever space syntax does, it seems to show how a complex system of relations can be a measurable empirical fact and, as such, constitute both an independent and dependent variable in the structure and functioning of a material system. This was brought to light by taking the highly improbable step of removing space from its embedding in the social and physical nexus of the real world and treating it as a thing in itself, as a pure set of relations. This idea would strike many as a clearly mistaken strategy, since everything that is interesting about space surely connects it both to society and to the material world. How can a theory of space be constructed by removing from space all that seems to make it relevant? It was only by extracting space from its embedding real-world context, and treating it as a thing in itself, that we were able to bring to light its configurational properties, and it turns out to be these that link space back to society, both as a receptor of social forces and also an active constructive agent in society.

The pathway from abstraction to measurable 'fact', in the study of relations in space syntax, is taken in two steps. First, the concept of relation is rewritten as the more complex concept of 'configuration'. Configurations are relations that take into account other relations. One immediate consequence of this is that a relation between two things that appears to stay the same can actually be configurationally different when embedded in a different relational context. For example, a pair of linked rooms off a corridor form a different configuration with the corridor depending on whether one or both are linked to the corridor, even though the relation between the two rooms appears to stay the same.[14] The difference between the two configurations, then, seems as hard-edged as we expect 'things' to be, in that each delimits possible human movement in a different way. In one case we must pass through one room to get to the other from the corridor; in the other we may go either way. Just as a 'thing' blocks our way, so a configuration of openings constructs possible ways and forbids others in a no less coercive, but essentially relational, sense. Configurations, which are constructions from relations, seem quite hard-edged things, even if relations are not.

The second step is that by correlating configurational analysis with, say, patterns of real movement, we can show that configuration has independent effects on the material world which cannot be mistaken for the effects of anything else.[15] In fact, if the theory of the 'movement economy'[16] is even partly the case, then it means that a complex relational entity – the configuration of the urban grid – drives the evolutionary dynamics of the urban system under the impact of social and economic forces. This argument does not depend on a cognitive connection by which we might argue that intelligent beings with immaterial minds have to decide where to move. If we move mindless agents in a computer randomly one step at a time in a configuration, then the agents will be distributed according to the connectivity of each element. Configurations, it seems, as composites made out relations, are empirical facts with predictable empirical effects – things in the real world – even if simple relations are not.

Strongly relational systems: society and space compared

The power of spatial configuration over what can be seen to happen in a city is so powerful that we are compelled to the conclusion that cities, seen spatially, are *strongly relational systems* (perhaps we should say 'strongly configurational systems'); that is, systems in which the relations of each element to all other are more important for the structure and functioning of the system than intrinsic or virtual properties of the elements themselves.[17] It is because they are strongly relational that spatial systems can be usefully conceptualised and analysed as very large graphs using configurational measures, which relate elements of the graph together however remote from each other they may be within the graph. The concept of a strongly relational system allows us to show that a set of related space-time events (such as the occurrence of particular spaces), which cannot be seen all at once, can nevertheless be shown to be real space-time systems with 'configurational' structures that are intrinsic to them, and that mediate their relations to other domains.

From the point of view of either a theory of society or a theory of society and space, this is an interesting formulation, since whatever else societies are, at one level, they seem to be relational – and perhaps configurational – constructs out of individuals. It is conceivable, perhaps, that the concept of a strongly relational system might be useful in understanding the superordinate relational systems that appear to be a key part of what societies are. Let us first consider some striking parallels between spatial and social systems:

- First, both are *dual* systems in the important sense that each seems to be made up of *material events* which take place in space-time, such as interactions and encounters, and *informational entities* such as the codes and conventions which seem to govern these material events locally (although not globally at the level of the emergent large-scale system). We might say that society has both hardware (interactions) and software (rules governing interaction).
- Second, both seem to be for the most part (in spite of utopias and ideal cities) *emergent* systems arising from *distributed* processes rather than designed systems. What needs to be explained, in both cases, is how an overall pattern of some kind is created over time by the independent activity of large numbers of agents in different locations.
- Third, both types of system seem to be *partially ordered*, in that each permits a great deal of randomness to co-exist alongside reproducible patterns.
- Fourth, both appear to be partly, or even largely, *nondiscursive* in that human beings operate, at least in their local patterns, competently without being able to say what it is that they are doing, so that while each is the outcome of human activity, and is utilised by human beings in everyday life, analytically speaking we have, at best, an unclear idea as to what it is that we understand.
- Fifth, both types of system seem, in spite of their *bottom-up construction*, to exhibit some degree of *top-down* as well as bottom-up *functionality*, in that, just as, say, movement and land-use patterns are functions of the overall structure of the urban grid, so individual social behaviours seem to be – although to a varying degree – functions of the overall pattern we call society.

These are striking analogies, on the surface at least. It is possible, then, that the concept of the strongly relational system might be useful in conceptualising what a society actually is. After all, we seem to be hard put to find any other way to conceptualise what it is that might exist above the level of individuals and constitute the 'real' counterpart of what we call society. On reflection, what space syntax has actually done, if it has done anything, is to bring to light space, as a hidden variable in the city, by showing its essentially configurational nature. Space was hidden, not because space cannot be seen (although this is also a problem) but because space is configurational, and configurations, like relations generally, are nondiscursive. We deal with them as ideas to think with rather than ideas to think of. Is it possible, then, that there are hidden configurational variables in society? And if there are, since they will show that space has been overcome and a large-scale entity called society created, can this also lead to ways to solving the problem of spatial necessity in society, and so pave the way to a theory of society and space?

Finding space in society

Where, then, should we look for a space-time strongly relational system in society? There appears to be only one possibility. What appears of society in space-time as its 'hardware' is interaction and co-presence, so we must look at these. However, if we look carefully at interaction and co-presence, then, at first, it seems to lead us away from the idea that society has significant space-time contents, and it does so for two reasons. First, although interactions occur in space-time they do not accumulate in space-time as spaces do to form a larger and larger connected system. They vanish, like blips on a computer screen, flashing on and off, and leaving no trace in space. Second, as soon as we observe interaction, we see that it is governed by conventions and rules that reflect who is interacting, how their statuses relate, what is going on, and so on. Thus, the same individual walking down the Tottenham Court Road will interact in one way with an old friend met in the street, another when ordering profiteroles, another having narrowly escaped a parking ticket and another having just acquired one. In interaction, the social software, rather than the empirical hardware, seems to be the important thing.

This invites an often-made comparison with language. The space-time events we experience are shaped by abstraction that we do not experience in the same way, yet it is the presence of these abstract rules and conventions in the space of interaction that render interactions intelligible as social events. As with language, although the rules are manifested only in individual behaviour, they must, in some sense, be independent of individuals if they are to carry the burden of making interaction socially intelligible. Since we cannot find these rules and conventions in space-time, we conclude that what is social about interaction are the abstractions that govern it. We reasonably conclude, then, that society, like language, is essentially an abstraction – one imposed on, and constructing, space-time reality but, in itself, an abstraction, nonetheless.

However, if we pursue the analogy with language a little more closely, we can begin to reconcile the idea of social rules as abstractions with a role for space-time in constructing society. As with society, we find it difficult to say where language is. We might say 'in the heads of individuals', but this raises problems. It is unlikely that any one individual has the whole of a language in his or her head and, in any case, what would happen after the demise of that individual? Also, language is pre-eminently a social thing, the property of a language community and constantly changing in some respects, reflecting its social nature, while remaining relatively fixed in others. However, if being in individual heads cannot account for all aspects of its existence, where then is language? One answer of course is that it is in space-time, in the dispersed language practices of the individuals who make up the language community. We may then say that language is reproduced through time by being realised in space. The 'language DNA' is out there in the real world of linguistic practices, not simply in our heads. The role of space-time in language is thus to be the medium through which language structure is reproduced by being produced.

This is not a bad role for space-time, and if it can be applied to the abstractions that govern language, then it can certainly be applied to society. A similar argument was used in *The social logic of space* to describe how cultural patterns of space were created and reproduced.[18] It was proposed that space acted an 'inverted genotype', in that the information needed to reproduce cultural patterns of space was to be found in the spatial configurations themselves (and it is this, of course, that we seek to retrieve through syntactic analysis) as well as in our heads. We then proposed that we interact with configurational information in the real world by our ability to retrieve abstract descriptions from concrete realities. For example, if one person builds a house and a second person places their house next to it, then a third person may 'get the idea' of a contiguous neighbour relation and place his or her house next to one of the existing two, and if this process continues, a line of houses will be created. Thus 'rules' guiding what happens in space-time, and leading to emergent patterns, do not have to exist in our heads as pre-programmed rules: they can be retrieved as logical properties from space-time reality, and used as templates for further action in it. Through this we sought to make our escape from the constrictions of the 'brain structure' theory of rule-governed activity, as put forward by leading proponents of structuralism.

A very similar idea, of course, underlies Giddens' concept of the *duality of structure* in human societies: that while being 'virtual', structure in society is both the medium and the outcome of 'situated practices' in space-time and these therefore link the production of social realities in space-time to the reproduction of their structures.[19] 'Structure' in society is, thus, comparable to that in language and can be conceptualised in the same way: it is both realised and reproduced in space-time. This is a compelling argument, although it does not deal theoretically with what may be a major difficulty that the 'inverted genotype' concept did attempt to engage: that societies are not simply embodiments of rules but emergent large-scale patterns that cannot be described fully through an account of the local rules that construct them. Most of what society is goes on 'out there', and although our knowledge of rules can generate emergent patterns through 'recursive activity' these do not include a description of the emergent structure. This seems to be a difficulty in principle with the Giddens' scheme. However, it seems to be a key fact of human societies that its members, at best, only poorly grasp the large-scale structure and, indeed, that may be why all societies seem to have specialists in retrieving descriptions of it. However, societies, like spatial systems, seem to create and reproduce their emergent structures largely by localised activity and, as we suggest below, this may turn out to be theoretically one of its critical properties.

At best, however, all these formulations identify space-time mechanisms by which structure is reproduced. They do not deal with structure itself, let alone the emergent structures that appear to come into being as much in society as in space. Yet it is exactly these emergent structures that we need to understand if we are to make sense of the large-scale changes in the spatial and institutional forms of society, such as those brought about historically by urbanisation and industrialisation, which must be accounted for by a theory of society and space. We need therefore a theory that is more than a mechanism for the reproduction of structure: we need to know how space is intrinsic, not only to how society reproduces its structures, but also to how society constitutes itself as a structure from local rule-governed activity. We must ask, then: is there any useful sense in which society is a space-time structure at the level of the emergent whole?

Systems of pure relatedness

Now, from the space syntax point of view, this is an interesting formulation, because space syntax, whatever else it is, has proved to be an effective method for retrieving descriptions of configurational structure from complex emergent realities. Every proposition that has been formulated about cities, for example, from natural movement to the duality of processes generating the grid, is rooted in this extraction of structure from complexity.[20] Since, then, cities and societies have so much in common theoretically, is it possible that there is also, in society structures, underlying complexity which might be discussed in the same way? This would seem to depend on how far it is useful to see society as being, or at least containing, in some useful sense, a strongly relational system.

What might we then be looking for? We can explore this by following Giddens' reasoning a little further. Giddens sees structure in social systems as 'virtual' because we find evidence for its existence only in dispersed practices, in the same way that we find evidence of language structure in discrete linguistic acts. This makes structure look rather weak, little more than rule-following. In fact, the cautious view of space-time, that leads Giddens to this view, seems unnecessary. The very idea of a society implies that, at some level, situated practices are likely to be connected. Although they appear as discrete events, none can exist in space-time isolation and no collection is likely to form a discrete system, not least because memberships of all situated practices are multiple and every individual passes continuously from one to another in a constant sequence. Each individual is therefore a link between a particular set of – for the most part recursive – situated practices, and all situated practices connect to each other through these changing memberships. Through the interconnection of situated practices, then, the individuals who take part in them construct a large graph of interaction, in which most individuals are remote from most others, but nevertheless have a finite depth from all others in the graph. Seen in a time-perspective, then, individuals are linked through participation in situated practices into a continuous system of time-space relatedness. On reflection, it seems likely that the existence of such a system is one of the preconditions for what we name as a society.

This 'system of pure relatedness' can, of course, be represented as a very large graph, in much the same way as a city is represented as a graph of its spaces. The problem is, of course, that although we cannot reasonably doubt the existence of the large graph, it is, to all intents and purposes, inconstructible. Even the most ambitious social network theorists, who use such graphs as a primary research instrument, only attempt to construct graphs for relatively finite and bounded socialities. What purpose can then be served by positing the existence of the graph of a whole society, when it is clearly an inconstructible entity? One possible justification is that, in spite of its inconstructibility, it is hard to doubt its theoretical importance. It is, after all, the global emergent product of the very situated practices that Giddens describes as the primary acts of social reproduction. If situated practices are the *means* of social reproduction, then the graph is surely its *product*, perhaps its only product. Its existence is a sign, and perhaps the only sure sign, that society exists as a system of interdependent, situated practices, linked by individuals, or interdependent individuals linked by situated practices. Once the large graph is admitted, it means that individuals are linked not only by abstraction or symbols but by practical space-time activity. On the basis of the large graph, we can reasonably claim that society is after all – or at least contains – a space-time system.

In any case, the fact that the graph is not constructible does not mean that we cannot know certain things about it. For example, we know that although the graph is very sparse, in that only a vanishingly small proportion of the potential connections between individuals are actually made, from the point of view of indirect connections, through intervening individuals, the graph is remarkably shallow. As Pool and Kochen show in their studies of finding graph paths from randomly selected individuals to others,[21] six steps are probably all you need, within national boundaries, and only one or two more across national boundaries. This is, theoretically, to be expected. If we think of those we know, and those whom they know, a beneficial combinatorial explosion from each step, out to the next, ensures that the graph from each individual to all others is remarkably shallow – and therefore highly integrated, in space syntax terms – in spite of being sparse.

Even this limited knowledge allows us to pose an interesting and highly general question (why should human societies be shallow graphs?) and find a simple evolutionary answer. For societies to be evolutionarily advantageous, all that has to happen is that those who are members of societies have to leave more surviving progeny than those who are not. How do societies do this? By spreading risk through the setting up of networks of interdependence. If my food supply runs out, someone else can help me. If something happens to me, someone else will care for my children. In evolutionary terms, a society is, at root, a network of interdependent relations that act as an insurance policy.

It is not too far-fetched, then, to suggest that the graph, or rather the network of relations that the graph represents, is what constitutes society, in the first place, as an evolutionary entity. If we then ask what social interaction is *for*, it is hard to avoid the answer: to construct the global graph, since evolution provides the rationale for its existence. Now, on purely theoretical grounds, we can assert that a highly integrated graph with, inevitably, a large number of cycles will be more robust than a tree-like graph since, in any tree, every time a link is broken, the graph falls into two disconnected sub-graphs. Ergo, a graph is likely to serve its evolutionary purpose to the extent that it is integrated. We note, as a corollary, that a large integrated graph will be better in evolutionary terms than a small one.

It follows that we do not need to provide further reasons as to why societies should seek to grow. It is taken care of by evolution.

Theoretical consequences of the large graph

Let us then admit the large graph as a legitimate theoretical, if unconstructable, entity. What else do we gain by positing its existence? The answer, we suggest, is that we completely change our view of what a society is by changing our view of it from a local to a global one. If society does, after all, have a global entity at its core, that is critical to its evolution, then it follows that critical, situated practices, through which the graph is created, will be selected for their ability to construct the large graph.

This is a serious cure for localism. It suggests that our fundamental, theoretical perspective on society should be at the global rather than the local level. One immediate benefit of this is that it allows us to make sense of social phenomena that had previously seemed puzzling. For example, if we look at the simplest, and least developed, societies known to anthropology, small-scale hunter-gatherer societies, we tend to find that their local groups, which many expected to exemplify 'elementary structures', are not only highly variable in size but extremely fluid in their composition. Individuals and small groups frequently take an excuse to leave one group and join another, usually where there is a relation of some kind to a typical woman. How can the extreme fluidity of local groups be reconciled to the sustained existence of a continuous and apparently strong society over a long period? Why is the weakness in the local groups not reflected in weakness in the society?

The answer is as simple as it is formal. Such societies, by definition, exist in environments where movement is required in order for individuals to survive, usually with severe limits on how many individuals can survive within a certain area. In such conditions, the large graph is much harder to construct than it would be, for example, in a group of dense villages quite close to each other. However, we can see that the social practices that lead individuals and groups to leave one group and join another will continually increase the density of connections in the large graph, if necessary at the cost of lower densities in the local graphs. The global graph of the society as a whole thus gains in strength, in spite of the weakening of the local groups.[22]

Similar mechanisms can even be found in less mobile – although not permanently settled in one place – societies. For example, Turner, in his remarkable study of the Ndembu, a village society that moves villages every few years, in which matriliny is combined with patrilocality, postulates a mechanism which similarly benefits the global society at the expense of local groups. Although the dominant ideology in the society – or at least among its males – is one of strong and large local groups under a local headman, in practice the majority of women (77 per cent), having lived with their husbands for long enough to have children, pick a quarrel and go back to live with their uterine sibling group, then 'after a period of ritual hostility' 'form joint hunting parties' with the husband's village. Turner argues that the high divorce rate is one of the fundamental mechanisms for strengthening the society as a whole, since the large-scale networks continually gain at the cost of weakening the local networks.[23] As with the hunter-gatherer societies, other institutional aspects of the society can be seen as supporting this prioritising of the global over the local.[24]

This is not, of course, the only way in which societies globalise. If we look at the Tallensi,[25] the society with which Turner contrasts the Ndembu so vividly, we find they live in scattered, but hierarchically structured, compounds in which women are spatially deep and men shallow (whereas in the Ndembu case, the buildings that form the circular village are simple huts). The Tallensi kinship groups remain in the same location through generations, and have deep attachments to their specific locations. In Tallensi society, the large-scale network of the society is created not by movement between groups (women in particular are relatively immobilised) but by a complex system of ritual, erected on the basis of an elaborate and hierarchical system of kinship, supported by an ancestor cult, and dominated by men. This forms an overarching structure that links relatively immobilised and localised groups of women, and it is largely realised through ritual interaction patterns that are confined to men. In this case, in contrast to the Ndembu, the integration of the large graph will be primarily through men, and largely through the realisation in space-time of highly ritualised and exclusive practices rather than through movement.

In each of these cases, we see that both the local and the institutional nature of the society is bound up with the social practices that create the global network, and is unintelligible without it. We also see that the effect of the institutionalised practices in each society is to create the global network, although in one case by strengthening (and rendering asymmetrical) the local networks and kin relationships and in the other by weakening them. In evolutionary terms, in both cases the institutional practices that have become selected seem to be exactly those that create the global network. We can say little about why one society takes one pathway and another takes an entirely different one, and it may be that we do not always have to do this to understand the morphology of a society. There may be specific historic causes, but it could equally be a matter of different responses to similar ecological conditions with some element of chance.

However, in evolutionary terms, a general mechanism may be suggested. If the large graph is created, in the first instance, by certain specific local practices, perhaps through a restricted random process of some kind, then, to the extent that the structure of the large graph is to be reproduced, the local practices which created it will need to be reproduced.[26] To the extent, then, that the system reproduces itself, the local practices will become normative in the system if for no other reason than that they are the means by which the large graph is reproduced.

This process would depend on the mechanism of description retrieval discussed earlier, that is the ability of the human being to retrieve an abstract description of spatio-temporal events and use it as a template for further action. Retrieved descriptions from practices, which had the effect of reproducing the emergent system, would, in effect, become normative to the degree to which the system was reproduced. There would seem to be no reason, in principle, to insist that the social practices that support the graph are antecedent to the graph. They may equally arise from the process of graph construction itself.

Suppose, then, that we tentatively define society as the large graph of pure relatedness plus all that it takes to produce and reproduce it, that is, all that it takes in terms of 'hardware' of situated practices and the social software that supports them. From the point of view of the society–space relation, this is interesting, because it means not only that a space-time entity exists at the heart of society but also that the existence of the graph means that space has been overcome to construct an entity at the level of the society itself, above the level of individuals and proximate groups. We have seen, in the few illustrative cases given above, that how society overcomes space to create the large graph may well be a useful clue to the morphological distinctiveness of that society.

We also see that the technology of production may relate to the ways in which societies overcome space by creating the initial spatial conditions – for example of aggregation or dispersion – in which space has to be overcome. Thus, a hunter-gatherer society has to overcome a degree of dispersion based on so many people per square kilometre, the Tallensi have to overcome place fixity of a rather dispersed kind, and so on. This suggests the possible form for the fundamental relation between technology, social institutions and space which have all been intertwined in the series of historic transformations of human society, in that the initial spatial conditions determined largely by the technology of production place constraints on the kinds of society that can develop through the interaction of space and institutions. Might there, indeed, be something like lawful pathways for the overcoming of space, linking the situated practices and social software to space? To answer this, paradoxically, we must first think of the graph in its relation to time.

Time and the large graph

Let us look a little more closely at possible mechanisms. Overcoming space means that a certain set of global relations in the network have to be created and realised in space, over and above those that arise from everyday productive activity. This implies movement. In some cases, as we have seen, this movement is created by some rather minimal social software permitting mobility between groups, in the form of rules that were both permissive, in that it did not require but allowed things to happen, and probabilistic, in that it was taken up opportunistically by a given, although substantial, number of people. In others, we find much more elaborate software that is both more constructive, in that it requires certain things to happen, and more restrictive, in that (as with the Tallensi) it specifies that men move but women do not. The global movement patterns that realise the large graph in space-time thus are far more rule-governed in some cases than in others. How do such differences relate to theoretical possibility?

Let us first consider the issue in principle. If we think of individuals scattered in a landscape, and of the movement required to create and reproduce graph relations, then we immediately see that there will be a simple law by which the probability of encounter is inversely proportional to distance. Other things being equal, I am much more likely to encounter those near me – or rather those near my daily path of 'effective spatiality' – than those farther away. This is interesting, since it suggests that for the global graph to be constructed, something like a natural law has to be overcome. Without some kind of social software, it seems likely that graphs would degenerate into localism. What kind of social software then can, in principle, operate to create a large graph?

We can explore this taking advantage of one of the fundamental concepts of space syntax: the justified graph, or j-graph. A core idea in space syntax is the concept of the graph whose elements are its j-graphs, that is, its positions from which the whole graph can be seen and be found to be different. The fact that there are such differences is not only the foundation of the idea of structure in space but also the basis of quantifying that structure: integration measures quantify what the graph looks like from the root of each of its j-graphs; representations of structure are colourings of the pattern of differences in j-graphs; and so on.

In societies, since the large graph links all individuals into a network that is eventually connected, it follows that each individual can be conceptualised as a j-graph of the large graph, that is, as the root of the justification of the graph from the point of view of that individual. One useful implication of this 'extrinsic' characterisation of individuals is that the individual and society are defined by exactly the same structure: an individual is a particular position from which the whole of the graph can be seen. Individual and society are then no longer polar concepts, but different ways of looking at the same thing. This also means that we can, to a useful extent, reason about the graph as a whole by reasoning about its j-graphs. Once again, it turns out that we can know useful things about j-graphs even if we cannot construct them, and find out how space gets out of the graph and into the social software.

Let us now return to Giddens. A society, he argues, reproduces itself by producing itself in space-time through rule-governed situated practices which thus, language-like, reproduce those rules. There is a corollary to this: *that what is not produced and reproduced in space-time is no longer an effective part of society*. There is, between the abstractions and the spatio-temporal events that make up society, a kind of *law of sufficient embodiment*: in the long run, *no co-presence*, *no relation*. This can be simply illustrated from our own experience of the way in which kinship relations decay, and no longer form part of effective networks, usually as a function of both real and logical space, in that, say, remote cousins who live in another part of the country are no longer an effective part of networks and become forgotten. If we do not retrieve a description of these relations and re-embody them in space-time encounters, then these relations are no longer a reproducible part of the social graph, although they may remain latent for a long while. On the other hand, an encounter that occurs unintentionally, that is without prior description retrieval (as is more likely in conditions of spatial proximity), will itself constitute a description retrieval event. The pair description-retrieval/interaction can occur in either order.

The relations that constitute a j-graph (and of course the whole graph) are, then, not only continually changing and being replaced but also being foregrounded by interaction and gradually fading into the background through inactivity. The periodicity of recursivity is critical. Every j-graph relation varies on a dimension of recursive realisation from frequent to never. The farther along to the never level, the more *virtual* the relation is, that is, a conceptual or potential relation rather than a 'real' one (although of course it may suddenly be realised again). Every j-graph then contains relations that go from real to virtual and each *time frame* will have a *sub-graph of real relations*, and the real *plus virtual* j-graph is the *total* j-graph.

As soon as we distinguish periodicity in time, a fundamental distinction that is found in all societies comes into view, which we can see by examining individual j-graphs. Each individual j-graph will reflect relations generated recursively through the *effective spatiality* of everyday existence, whether the individual is a hunter-gatherer or an office worker commuting from Chalfont St Giles to High Holborn. These relations in the graph, weak and strong, will be generated as a by-product of co-presence generated by how individuals produce their everyday survival through productive activity and its associated technology. Second, the graph will *reflect relations that are generated by the need to reproduce the graph itself*. This may sound odd at first, but is a key distinction. In all societies, a distinction can be drawn between activities whose object is the biological survival of its members, and activities whose object is the production and reproduction of society, for example, the special activities associated with major life events such as birth, coming of age, marriage and death. *In the first case, the graph is the by-product of the activity, in the second, the graph is the object of* the activity *and so* is, in a sense, *its product*. This distinction corresponds to what economic anthropologists have called the 'replacement' and 'ceremonial' funds,[27] the first being the proportion of human resources and effort devoted to reproducing the ability for individuals to survive biologically, and the second the proportion of resources devoted to biologically useless activities whose object is to produce and reproduce society.

For a theory of space and society, this distinction cannot be too highly emphasised, since it is involved in every phase of how societies create and recreate themselves in space, and thus overcome space. In the first kind of activity, which covers the necessarily spatial conditions of everyday life in which individuals engage in 'work', relations in the graph are *generated* through an activity which has another purpose. The graph arises because work happens and creates interactions that may or may not be reproduced within the work process. Everyday activity, aimed at everyday purposes, can then be said to be *generative* of the graph. In the second kind of activity, which covers the special activities in which individuals engage which do not have biological survival as the direct aim, and in which the graph itself is the object of attention, relations in the graph are *conserved* through realisations that are designed to achieve just that. Such activities are thus *conservative* of the graph. This is what we mean by ceremonial activity, whether simply inviting people to dinner or engaging in some large-scale social ritual: its aim is to reinforce and reproduce the graph through activities in which the content of description retrieval of relations in the graph is manifest and dominant. This is what we mean by, say, a marriage ceremony or a coming of age ritual.

Ceremonial, graph-conserving activities are, thus, distinguishable from everyday, graph-generating activities by their degree of deliberate, descriptive content. The content of a ceremonial, or ritual, activity is subject to a greater multiplicity of rules governing what happens, including who does what and in what sequence that will be found in everyday life. This is in its nature, since its object is to describe with great emphasis and re-embody in space-time key relationships in the graph. We can formalise this by working out the number of rules that are required to make the event happen against the number of activities that actually happen, a kind of rules-over-events ratio. The higher the ratio of rules to events, the more we would say that the activity is 'ritualised'. We can therefore say that the description required for a ritual is a long one, in extreme cases as long as the number of events taking place, since nothing can happen unless it is specified by a rule. We can conveniently call such activities 'long models' in that they depend on a long description to ensure that they happen in the proper way. In contrast, we can immediately see that everyday activity tends to be 'short model', in that insofar as its objectives are practical rather than ritualistic, it will only be effective to the degree that the actor is free to carry out the necessary activities in an unencumbered way as possible.[28] We thus find a fundamental relation between time and social software. In normal circumstances, short model, graph-generating events have a short periodicity, and long model graph-conserving events have a long periodicity. A fundamental dimension of difference in what we might, with some licence, call universal social software is, therefore, bound up with time.

How space gets into the social software

However, it is even more powerfully linked to space through the fact that time is linked to space through movement. All movement occupies time, and all encounters depend on some degree of movement. The question is: what degree? In terms of the relation between movement and encounter, there is a profound difference between local (in terms of the effective spatial pathways of individuals) and global space. In local space, encounters happen through the agency of space itself, and such encounters can either produce new relations in the graph or reproduce existing ones. In local space, encounters happen with little effort, and there is no reason for anything more than the minimum of rule structures to bring them about. If we find complex rule structures, or long models, in local space, as in the case of the Tallensi compound, then they are likely to be about restricting encounter rather than generating it. In its natural state, local encounter is a short model.

Encounter over distances is quite different. At the very least, a distant encounter will normally need to be aimed at a specific object, a destination that must be specified in advance. Precisely because the probability of encounter is inversely proportional to distance, a distance encounter needs a greater degree of conceptual organisation than a local encounter. As usual, we find that these ideas still pervade our unconscious assumptions about everyday life. For example, we assume that an impromptu invitation is much more easily issued to someone who is local than someone who must travel a distance, and that more formality is required if a greater distance is to be overcome to make the encounter happen. We thus find that another invariant aspect of social software is that, complementing the inverse laws linking encounter to distance, we find another that *links distance directly to the length of model*. Events that are rarer in time are also, normally, more distant in space. We therefore need the greater conceptual organisation of the longer model to bring about encounter over greater distances, and this increases with distance.

However, there are two kinds of distance: real distance and social distance. Longer models are found where either is to be overcome. For example, a person in the local domain whose status is asymmetrical to others will have long models associated with interaction, since the need is for social software that restricts and structures encounter in a domain where it may otherwise happen by chance and in an unstructured way.

This is fundamental to the ways that certain kinds of buildings work: space is structured to conserve formal relations by preventing spatial proximity turning them into informal ones. We can now see a natural morphological logic in the comparison between the Tallensi and the Ndembu. Among the Ndembu, we find greater equality between men and women and a less asymmetrically structured society, because the large graph is prioritised over the local graph and is created by the short model movement of women away from the husbands and back to their uterine sibling group. This solution to the large graph problem, by movement, means that the local short model is pushed outwards to form the larger network. Among the Tallensi, the contrary is the case. The large graph is created, not by short model movement, but by long model movement-based (the shrines are remote from compounds) ritual activity of men, based on the lengthening of the model of the kinship system (descriptions of lineages are retrieved much farther back among the Tallensi than the Ndembu) which excludes women, who have more localised, and also much more structured, spatial lives. The Tallensi therefore can be seen as extending the long model downwards from the global to the local level. The Tallensi–Ndembu contrast thus becomes a pair of morphological opposites, extreme cases of two ways of creating the larger graph.

This is how space gets into the social software. The key items of social software are about space as well as time: on the one hand, longer models are needed to overcome space to create the global graph; on the other, they are needed to control the effects on graphs of spatial proximity. In other words, we need long models to overcome distance and to controlling proximity in the reproduction of graphs, and we need short models to generate and sustain graphs in the first place.

Urban societies

What then happens when these initial spatial conditions, that create the kinds of societies we have so far considered, no longer prevail, for example, when societies aggregate to form the large continuous and dense settlements that eventually become cities? Why do spatial and social changes seem to happen together? How do both relate to changes in the technology of production: to quote Wirth, 'The central problem of the sociologist of the city is to discover the forms of social action and organisation that typically emerge in relatively permanent compact settlements of large numbers of heterogeneous individuals.'[29]

Let us first add a problem. There is something approaching a paradox in our historic idea of the city. It seems to be, at once, the locus of domination, social classes, bureaucracies and enforceable law. On the other, throughout history it has seemed to be the place of freedom. Is it possible that both of these effects are, at once, a product of what the city is? As previously, we will not try to explain either the changes in the technology of production that are associated with cities or why, and how, it creates specific spatial conditions. We will take the spatial condition of dense aggregation for granted and try to understand its implications for the construction of large graphs.

First, from the previous paper in this issue (and its predecessors) we already know a good deal about the impact of space in large dense aggregates: a varied – but sometimes intense – pattern of co-presence is generated through the effect of spatial configuration on movement. How is the large graph likely to be affected by these very different initial spatial conditions? We saw that, under dispersed conditions, the problem of creating the large graph was a problem of overcoming the distances between relatively small local groups, that is, it was largely a problem of creating co-presence through movement in spite of dispersal. What, then, is the effect of the replacement of those initial conditions by the conditions in which co-presence is much more freely available?

First, let us consider the institutional changes commonly associated with the rise of cities against this change in the background conditions. A vast literature suggests, with reasonable consensus in spite of exceptional cases, at least three main changes. First, a substantial division of labour among individuals appears to replace small-group self-sufficiency. Second, space-based supra-local organisations with a predominantly political character, and the ability to settle disputes according to agreed law, take the place of supra-local organisations based on the elaboration of kinship structures that lack significant dispute settlement functions and are articulated largely through ceremony. Third, and this is much harder to define, the notion of the psychologically free individual takes the place of the highly constrained, social member whose group identity is of greater social importance than individuality. Why should these be the outcomes of the transformation of space?

The first outcome, the emergence of an individual division of labour, is not difficult to relate to the spatial transformation. An extended division of labour among individuals, and the intricate pattern of day-to-day interdependencies that this creates, is inconceivable without the integration of space and the high levels of natural co-presence it makes possible. We cannot say that the division of labour is caused by spatial integration, but we can say that if there are economic or evolutionary reasons, in which this division of labour is advantageous, configurational integration creates the necessary spatial conditions in which it becomes viable. A division of labour among individuals becomes ineffective to the degree that distances between specialists become greater.

The rise of space-based political institutions is also closely connected to the transformation of the spatial basis of society, although the relation is less obvious. In preurban societies the task of supra-local organisations was to overcome the distances between the groups and create the larger-scale society out of spatially dispersed groups. The raw material for this was the kinship system that already creates relations across space, usually supported by the exchange of people between groups through marriage and other alliance-creating acts. Supra-local organisations in these spatial circumstances tend to raise the kinship system to a higher level and embed this back in the society through ceremonial organisation. This is why dispersed, tribal societies often have higher levels of ceremonial organisation than urban societies, with a greater presence of supra-local ceremony in everyday life.

Under the spatially integrated conditions arising in an urban society, the problem for supra-local organisations is different. We have already seen that space gets into the social software by creating rules, on the one hand, to overcome distance and on the other to control the effects of spatial proximity. In preurban societies, the first of these is much more important than the second, since without it the global graph would not exist, and the second arises as a reflection of the modality of the first. In densely aggregated societies, the second takes priority over the first, not only because the compression of space has made larger graph resources available much closer, but also because the problem of controlling the effects of proximity has become more important than the problem of overcoming distance. In dispersed societies, when disputes occur, the common solution is either fission[30] or the formation of kin-based alliances to redress the situation. Under urban compression, the potential for disputes is statistically much greater and, by definition, the fission solution is no longer available if the dense society is to evolve.

In practical terms, it means that the ability to settle disputes, within the spatial realm, and prevent the graph from breaking up has become more important than the need to construct the graph across distances. In terms of the language of description retrieval we can say that the need to control and negotiate descriptions in the continuous, spatial milieu has become more important than the need to embed descriptions in ritual in order to create the large graph. A 'political' organisation is one that specialises in the negotiation of descriptions, and a 'legal' organisation is one that specialises in the control of descriptions. It is such organisations that are then selected in the new spatial conditions created by dense aggregation. The problem of distant relations has not, however, disappeared, but reappeared in the form of the need to relate a much larger aggregate to the wider system, including the urban hinterland and the other settlements in the wider system of which the city is part.

These factors impart to supra-local organisations a character that is not only political and legal but also space-based, both in the sense that it must operate within a spatially continuous local system, and also in the sense that it must relate this system to other spatially based systems in its vicinity. We can say then that, whereas, under dispersed conditions, supra-local organisations create society in spite of the lack of spatial integration and, therefore, use primarily ceremonial means, under urban conditions they create society in spite of the presence of spatial integration by dealing with the problems it creates, using primarily political-legal and space-based means.

It is the change in the institutional structure of society, in response to spatial changes, that then creates the third phenomenon: the emergence of the psychologically free individual. This is normally assigned to the dramatically increased co-presence resulting from spatial integration changing the everyday experience of others from social recognition to anonymity, and discussed in terms of the flawed discourse of the 'myth of historic spatiality' as some form of alienation or desocialisation. We propose a deeper cause, arising from changes in the interpenetration of the spatial and supra-local aspects of society. In a pre-urban society, social institutions work on the raising of a kinship network into a larger-scale conceptual system, so that these become the most important aspects of the global organisation of society. This means that the burden of reproducing the large graph is carried largely through what people do, how they think, and how they behave. The load of reproducing institutional structures is carried through individual minds and individual behaviours (although forming collective patterns) and, in this sense, the individual in a pre-urban society is much more a mental prisoner of that society.

In the spatial conditions created by the city, we find not only that institutional structures have become transformed but also that they have become spatialised in two senses: they are located as built forms in real space, usually in significant locations in the urban fabric, and in the sense that their sphere of influence is now the ambient space itself, not simply the abstract realm of social software that served to create the momentary space-time events that reproduce the large-scale structures of the graph in pre-urban societies. Social institutions are, in both senses, taken out of the fabric of individual life and made extrasomatic. Institutional structures, in effect, are externalised from people and become an outward pressure bearing down on them, rather than an internal force that structures their thought and behaviour. Although they act as external forms of control operating on the individual through the control of ambient space, they also liberate the consciousness of the individual, and turn him or her more fully into a social individual more than simply a social member.

It is not, of course, that individuals do not occur in pre-urban societies. One remembers the remarkable, and highly spatial, commentary of Lienhardt on the Dinka-Nuer and the Anuak:

> The frequent dispersion of the Dinka-Nuer as compared with the concentration of the Anuak, may be associated with the much greater interest shown by the Anuak in individuals and personalities. They have an extensive psychological vocabulary, and their village politics are conducted through an interplay of character as well as of faction. Anuak are interested in people, Dinka-Nuer more interested in cattle.[31]

However, in the city, the creation of extrasomatic institutions frees the mind from the need to use its network simply to reproduce the existing structures of the graph, and sets the scene for non-local networks based on a choice of the kind that Fischer (see below) describes as characteristic of present-day cities, but which have, probably, always been one of their prime assets. The extra-somatisation and extra-mentalisation of institutional structure within a context of intensive spatialisation is the prime source of the individuality that seems to be associated with cities, and we can note that it arises from the same spatialisation of social forces through which class asymmetries and bureaucracies also arise. The city, itself, becomes the extrasomatic mind, and this frees the internal mind and makes it creative.

What then happens to the large graph and its constituent set of j-graphs in the context of the city? Let us look more closely at the studies of Fischer et al. on social networks in contemporary cities.[32] As a result of their investigation of networks, Fischer et al. are highly critical of the tendency of previous investigators to focus on the local properties of networks, that is, such properties as density (if a knows b and a knows c, then b knows c) and multiplexity (if a knows b because he is his brother-in-law, does he also know him as his butcher?). They see this as part of what we have called the 'myth of historic spatiality', in which the present is believed to be alienating, because individuals lack embedding in dense and multiplex local networks. Fischer's view is that if people had predominantly local networks in the past, then it was because they could not escape from them. Compared with this enforced localism, networks in the modern metropolis were of higher quality, more dependable and perhaps also more extensive, because they were formed by choice and affection rather than dependence on locality. Fischer and colleagues' work is one example of a growing group of studies that suggest that the more global – or perhaps globalising – aspects of networks may be more critical than the local.

We cannot, of course, get social network information on historic cities, but we can reasonably conjecture certain, likely properties of j-graphs from a knowledge of living patterns and institutional structures. For example, if we take a late, medieval, mercantile city, like London, a typical individual would be a member of numerous different networks which do not correspond to each other and may barely overlap. For example, during the working day the citizen would, through his or her part in the division of labour, be part of the network of making, distributing and trading, that is, what Durkheim called an 'organic solidarity'. Also, through the division of labour, he or she would be a member of a guild, which would make links into a quite different network, one more like a Durkheimian 'mechanical solidarity.' In all likelihood he or she would also be a member of a religious grouping, and of a family network, which again would make links to different groups. Far from being multiplex and dense, the citizen's j-graph seems to be based on multiple overlapping memberships that have the opposite properties. Such graphs may surely be seen as globalising rather than localising. At the same time, we would expect that the mechanisms for regulating locally co-present relations would be reinforced, giving rise to the familiar urban theme that socially successful people rarely network with their neighbours.

The typical urbanite is, then, one who globalises networks. Dependence on purely local networks is surely correctly taken to be a sign of under-privilege and lack of social power while, at the other extreme, those with the most global networks are also likely to be those with most social power. We can see how cities will tend to create the full range of local and global networks. However, in general, we can see that *cities are machines for globalising networks through multiple memberships*, and regulating local networks to cope with the strains of co-presence. The urbanite is successful to the degree that he, or she, succeeds in the globalising game. In the last analysis, this all seems to be the morphological consequence of the integration of space, and again we are reminded that it is the globalising, rather than the localising, aspects of social software that are critical to understanding the society–space relation.

In many senses, then, the social nature of cities seems to arise in quite a natural way, not only from the fact of aggregation, but also from the form of aggregation. The social city would be inconceivable without the fundamental network of linear spaces that link all parts of the city into a unified and structured network of movement and co-presence.

So, what is happening now?

So, what can be learned about cities now? First, let us review our theoretical lessons about cities:

- although cities, as systems of built forms and spaces, are driven by economic and social processes, they are not infinitely plastic in the forms they take, but evolve under the constraint of spatial laws governing both the emergence of spatial forms and the effect of these on co-presence;
- the relation between urban life and the city, as object, is as a result generic not specific, reflecting the generic nature of the relation between space and society rather than the idiosyncrasies of history;
- space, and most notably city space, does not just take its shape from the society, but answers back and affects society, even changing its deepest structures.

and about societies:

- all societies are *global* (although obviously not on the same scale) in that the global graph is the spatio-temporal sign that a society exists, and *virtual* in that they operate on social software of different kinds to create their graphs in different spatial circumstances;
- a society is how it overcomes space to create its large graph, and the tendency of individual j-graphs is toward the global graph, at least as much, if not more, than to the consolidation of the local graph;
- the effect of the technology of production, including the patterns of effective spatiality that it requires, is to create the initial spatial conditions in which society, the large graph, and the software and hardware through which it is realised can be created.

What then do we learn about the here and now, unencumbered by the myth of historic spatiality? First, we can define certain questions that the theory suggests might be critical:

- What are the new spatial preconditions on society imposed by new forms of production?
- How will they impact on the two-way genericity of the city?
- Will the distinctively urban dynamics that we have described continue to prevail, or be replaced by others?

First, let us review the effect of industrialisation on cities. It clearly created, through a new technology of production, new spatial conditions in which a society had to be created, at least for some people. In practical terms, the factory system meant that the rapid agglomeration of a large number of factory workers, in new and rapidly growing urban agglomerations (some new, some extensions of existing ones), brought a large number of people into the city who were cut off from their previous, social embedding, and did not obviously fit into the social and spatial patterns of the preindustrial city not least because, under the factory system, the artisan was separated from the tools of production, and no longer had the material basis for the urban memberships we have described.

One outcome of this was an excellent example of our model in action: social thinkers saw that a new society had to be created for the people brought together under these new spatial conditions, and a series of fantasies and experiments were proposed (and many carried out) for spatially redesigning society in order to achieve a pacified social system.[33] These fantasies, which all involved the disaggregation and dispersal of the city into smaller self-contained components, were deeply influential in the intellectual origins of modern town planning and were a key factor in creating the over-localised thinking that has prevailed since then in the spatial disciplines.[34]

Other outcomes were equally well known. We saw the auto-generation of space-based, urban communities of the kind reported for the East End of London and the West End of Boston and now duplicated in informal settlements in cities around the world.[35] Large numbers of middle-class people, who had previously inhabited the more central or inner suburban areas of the city, moved to the outer suburbs and the countryside, and a programme of spatially controlling the poor, urban population was initiated, first by building closed urban enclaves in city centres in place of streets,[36] and then seeking to 'thin out' the urban population (that is, get rid of the poor) by 'decanting' them into new dormitory towns (including, eventually, the English New Towns programme). In less-advanced, industrial countries, this was complemented by a programme of re-engineering the grids of city areas to make them fit for largely middle-class populations (Haussmannisation, Cerdaisation and so on).[37]

It is of theoretical interest that this programme of deliberate, urban disaggregation and amelioration was perhaps the first time that a programme of creating a society, under new spatial conditions imposed by technological development, was a conscious, discursive process. No such fantasies are proposed for whatever change in spatial conditions is currently underway and this may, itself, turn out to be an interesting fact. But what is the current process? First, at a descriptive level, we are seeing a dual movement in, and out, of cities. People are still leaving cities, as they have done for many decades, but others are coming back. Most cities now report quite a rapid recolonisation of old city centres, for the most part by those seeking a street-based lifestyle – gentrification is essentially a house-on-a-street process, and usually back to houses originally intended for people like them, but which had, in a previous era, become de-gentrified – but also by the construction of high-priced inner-city enclaves in some areas. Many cities, and not just in the Western world, also report a rapid intensification of street life in cities, in many instances reversing a long-standing trend the other way.

To say the least, this sits uneasily alongside the theories that propose that modern communications technology means the end of cities as we know them. What we are seeing seems to be a straightforward revival of urban living, in many senses, apparently, simply reversing the outflow that took place over the last century. If this does turn out to be the case, how could it be explained? Who is going where and why? And will it fit into cities more or less as they are?

The answer must start from trying to understand the change in the 'initial spatial conditions' that are now being imposed by new technology of production. What exactly is the change? At the risk of serious over-simplification, could we not suggest that what is happening now is something like an inversion of the change that happened under industrialisation? Then, the changing technology of production brought into cities large numbers of desocialised, de-tooled and disadvantaged people to man the new systems of production. The flight from the city was of a smaller number of traditional urbanites (so that cities still grew rapidly overall), and presumably this was associated with the de-gentrification of some urban areas that are now re-gentrifying. Now, we are surely seeing something like the opposite: the dispersal of low-level production and services away from major conurbations, and the aggregation of high-end knowledge production in and around cities, especially their centres. It would be expected that, under these initial spatial conditions, there would be a two-way flow both into, and out of, cities.

But what exactly is driving the flows? First, the leading-edge change is in knowledge-led industries, and the new activity types that are being created are to do with the creation or transformation of knowledge in some sense. The dispersed, productive activity is new in its contents, but it is not a new type of activity. Under industrialisation, the leading-edge change in activity was not in the knowledge that led to new forms of production, but in the material fact of production itself. It was this that created the spatial conditions in which society had to operate. We could say that, compared to the nineteenth century, the leading edge of activity change has moved from the hardware to the software of production.

Why should this lead to an inflow into cities? There is a work-related and residential aspect to this argument. The first is that knowledge-generating industries – and to a lesser extent knowledge-rich industries – are buying into integrated space for reasons that are as profound as those that first associated the progressive division of labour in cities with the integration of urban space. It is more than 20 years since Tom Allen of MIT showed that the intensity of contact between R&D groups was positively related to innovative performance, but intensity of contact within groups was not.[38] It is about 10 years since space syntax research showed that increasing useful contacts between groups resulted from the way in which layout configuration related global movement to local working patterns.[39]

The reasons for this apparent relation between space and innovation are simple enough: in innovative knowledge generation you are more likely to make unexpected links by talking to those working on other problems than by continuing to talk to the same people. The potential of a new idea is greater if it comes from a greater conceptual distance and this is more likely if you talk to people who do not share your problems and your take on things. Space can create the conditions in which this becomes more probable by making it more likely that you will talk to people you do not know you need to talk to. This is why innovative thinking benefits from a relatively random and rich background of encounters and suffers in an over-organised or hierarchical one. It is not the relations within the group you work with that need to be reinforced, but relations between groups. Again, it is an argument for globalising, rather than localising, graphs. In this it perhaps resembles the original development of the division of labour in cities. This similarly depended on something like critical mass in the graphs at a non-local level. Indeed, it is likely that, as Brown has suggested, the integration of knowledge-generating space itself both promotes and benefits from a progressive development of an interdependent division of knowledge-related labour.[40] The grounds for seeking the integration of space for knowledge-creating and knowledge-rich activities are therefore as profound as those that originally took place in historic cities.

But why should this also lead to a resurgence in urban living? There could be an equally simple answer: for the traditional urban reason that cities are not about creating localised networks but about globalising networks. More precisely, urban living allows, as it always did, the putting together of the middle- to long-range networks that are an indispensable adjunct to the knowledge-creating society that grows up alongside the knowledge-creating process, for the kinds of reasons shown by the classic studies of Granovetter in which he showed that it is not the immediate network of strong ties that give the critical information about work opportunities, but the looser network of what he calls 'weak ties'.[41] Again, the larger-scale network is more critical to the individual than the local network. Put at its simplest, if you are an upwardly mobile part of the growing knowledge industry, you can hardly afford to be out of the city. The risks to your j-graph would be too great. The city still operates as a machine for globalising graphs. Why then are others still leaving? Again, there could be a simple answer: because the leavers are not part of the leading edge of the new technology or the new society, and so are still working under the old paradigm of escape to the suburbs and the countryside. They have not yet become part of the new society.

How will this new pattern then fit, or fail to fit, into the existing city? We have specified a micro-economic process that requires integration, and a residential process that requires some degree of integration. As we have seen, the process of 'gentrification' the world over is led by the traditional 'house on the street', although in difficult areas for 'urban pioneering' we still see the implantation of enclaves, perhaps to facilitate the process in the short term. In other words, the new kinds of activity, generated by the new spatial conditions of production, seem to be calling into existence something very like the traditional city as we described at the start of this paper. In other words, the criteria we suggested earlier for the absorption of new activities into an existing framework seem to be fulfilled, and this seems to be confirmed in the case study reported in the Appendix. This can hardly be the end of cities as we know them. It is surely another new beginning.

Appendix: Case study in fast and slow change

This is a simple example, demonstrating both how the fast rate of change in urban activity patterns is absorbed by a slow rate of change in spatial form, and how the existing spatial pattern of the city acts constructively as a generator in bringing about new patterns of spatial culture within a largely unchanged urban fabric. The case is the historical transformation of the City of London over three hundred years from an urban community based on guilds and face-to-face exchange mediated by an intricate and dense system of streets and alleys, to the current financial centre where most business is conducted by computer and behind closed doors.[42] How has the city responded to this transformation of its spatial culture, and what, if any, is the effect back on this life?

Figure 18.1 is a black-on-white representation of the space of the City of London as it was in 1676 shortly after the great fire, and Figure 18.2 is its axial map. Figures 18.3 and 18.4 are the same for the City in 1989. Let us first look at the changes to the urban space, leaving aside at this stage the dramatic increase in building heights and densities in certain parts. The differences are quite systematic and diffused throughout the system. For example, the number of convexly distinct spaces has reduced to about half, as has the number of axial lines. The number of dead-end lines has reduced to about a tenth of what it was, but the number of through-lines has also reduced. Also, the number of built form 'islands' has

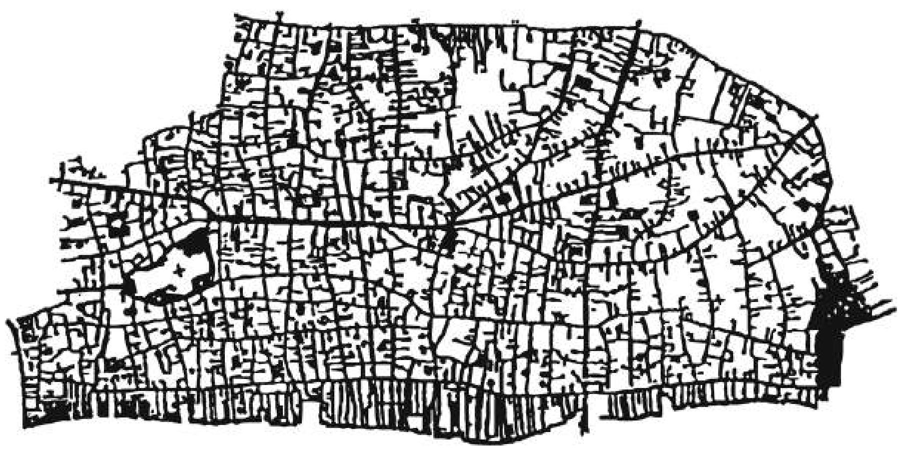


Figure 18.1 Space in the City of London in 1676.

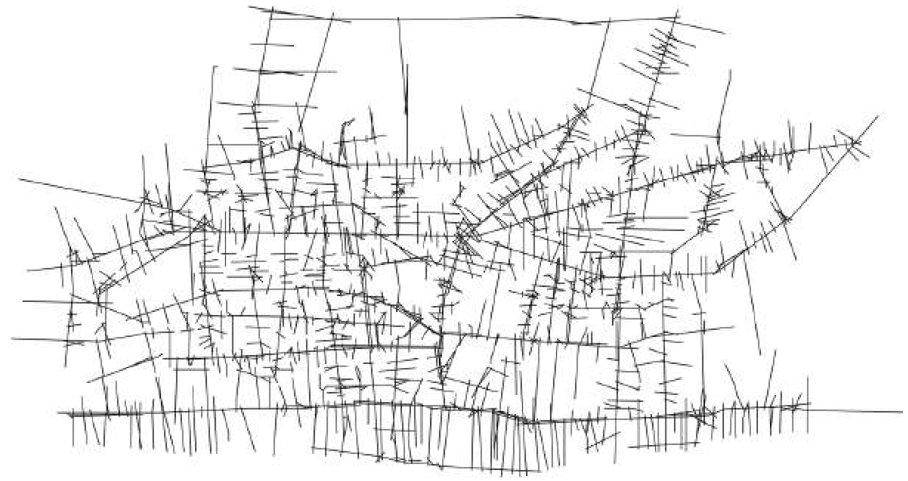


Figure 18.2 City of London 1676: axial map.

Figure 18.3 Space in the City of London in 1989.

reduced by about a quarter. Geometrically, the spaces have become considerably wider, and lines are on average substantially longer. There are three major new streets – England's answers to Haussmann in the mid-nineteenth century – forming a triangle meeting at the syntactic centre. Overall, the system has become much more syntactically integrated, and more intelligible, and, in spite of the overall reduction in the number of blocks, more permeable.

We also find that the 'two-step logic' of the city is conserved and even improved. This means that at the global level, that is if you enter the City by one of its gates and follow at each stage the longest line you can see, then you will see the centre

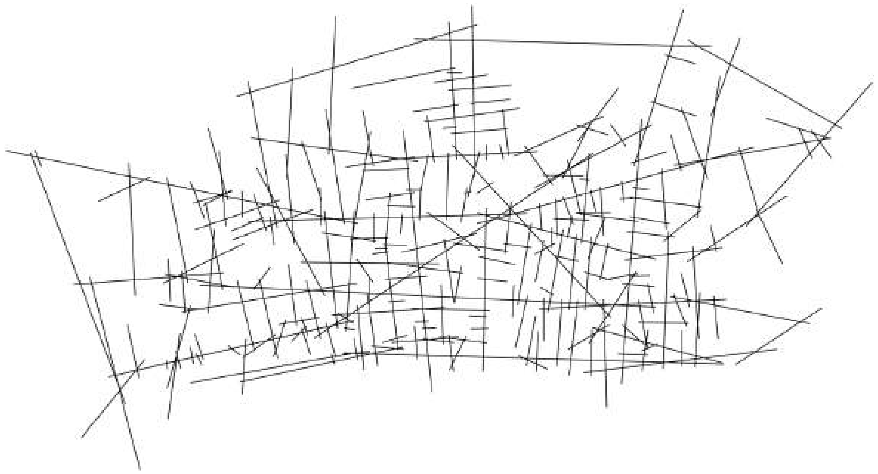


Figure 18.4 City of London 1989: axial map.

somewhere from the second line. At the local level, it means that if you depart from a main line into a shorter line into a back area, then the second line you use will show you either the way out or an important internal destination. It means, quite simply, that it is hard to get lost because, both at the global level of the whole City and at the local level of its small-scale subareas, the City has the kind of centre-to-edge structure that we saw in the previous chapter. The effect of this is that the City works to create strong probabilistic interfaces between those moving in and out of the buildings and those moving past, and between those moving in the larger-scale system and those moving in the smaller-scale local areas. The City constructs, in effect, a series of probabilistic interfaces between scales of movement, so much so that we can be sure that it has evolved in order to create this kind of pattern of co-presence. The need to interface scales of movement is the situational constraint that has governed the process by which the urban pattern has emerged.

It is easy to imagine how this local, and global, space structure would have supported the face-to-face commercial community that occupied the City three hundred years ago. But how does it work now? The fact is that it works in more or less the same way, but with quite a different social embedding. One has only to spend time in the City of London to realise that it has an extraordinary spatial culture. Those who work in the City go out into the streets, especially in the midday period, and use public space for eating, drinking and socialising. In the recent past, this has been substantially added to by the building of highly successful new public squares, whose use now often extends well into the evening. In some cases, new experimental designs of public space offer new kinds of urban experience by engineering new kinds of co-presence. For example, Broadgate's Exchange Square creates a number of focal spaces within the same large space so that the different groups which congregate in its various parts are all in visual contact with each other. The effect is exhilarating, and generates a substantial amount of interaction through this engineered co-presence.

How should we then understand this? The system of public space is still being used in generative mode, but not so much as a direct support for face-to-face business activity, but to create an emergent spatial culture which in every sense stands in contrast to and thus complements the business activity. For example, while business activity behind closed doors is oriented towards gain, the sociality of the public spaces is oriented to gift exchange and conspicuous consumption. The situational constraints that hold this evolving system of public space in place are to do with the construction of a complementary sociality, to bring together in society what the business activity divides and brings into conflict. Space thus plays as powerful a social role as it ever did, but in a different modality. It constructs an expressive rather than instrumental sociality. More practically, the distinctive 'spatial culture' of the city is a prime component of the famous quality of life that draws both individuals and organisations to the city. The new sociality has economic consequences.

Why has this new lived sociality emerged? It seems unavoidable that part of the answer is simply that the spatial and co-presence preconditions which it requires, and which can generate the co-presence conditions in which it can emerge, were already in place. Spatial culture is endlessly created and recreated by the spatial and institutional conditions that we impose on ourselves. It is a perpetually emergent phenomenon, arising from the fact of co-presence and the fact of society. What we have to understand is how it is structured, and this can only be in terms of how spatial conditions provide the co-presence raw material within which different spatial cultures will emerge. The city itself creates its spatial culture.

Notes

1. J.-F. Lyotard, *The postmodern condition: A report on knowledge*, translated from the French by Geoff Bennington and Brian Massumi (Manchester: Manchester University Press, 1984).
2. S. Sassen, *The global city* (Princeton, NJ: Princeton University Press, 1991).
3. D. Martindale, Prefatory remarks to M. Weber, 'The City', in *The theory of social and economic organisation*, Eng. edn (Glencoe, IL: Free Press, 1958).
4. P. G. Hall, *Cities in civilization* (New York: Pantheon Books, 1998), p. 943.
5. This article is a revised version of a paper with the same title that was delivered to the *Third International Space Syntax Symposium*, Atlanta, GA, 2001 (eds).
6. B. Hillier, *Space is the machine: A configurational theory of architecture* (Cambridge, Cambridge University Press, 1996).
7. J. Hanson and B. Hillier, 'City of London: Shape and role', *Urban Design Quarterly* 48 (1993): 8–13.
8. See, for example, R. M. Adams, *The evolution of urban society: Early Mesopotamia and Prehispanic Mexico* (London: Weidenfeld & Nicolson, 1966).
9. See, on hunter-gatherer societies, R. B. Lee and I. Devore, *Man the hunter* (Chicago: Aldine, 1968); on tribal societies, V. Turner, *Schism and continuity in an African society* (Manchester: Manchester University Press, 1957); on social networks, M. Granovetter, 'The strength of weak ties: A network theory revisited', in *Social structure and network analysis*, edited by P. Marsden and N. Lin (Beverly Hills, CA: Sage, 1982), pp. 201–233, C. S. Fischer, R. M. Jackson, C. A. Stueve, K. Gerson, L. McCallister Jones and M. Baldassare, *Networks and places: Social relations in the urban setting* (New York: Free Press, 1977), I. de Sola Pool and M. Kochen, 'Contacts and influence', *Social networks* 1 (1978): 5–51, reviewed in G. L. Albrecht, R. Fitzpatrick and S. C. Scrimshaw (eds), *Handbook of social studies in*

health and medicine (London: Sage, 2000); and on organisational dynamics, T. Allen, *Managing the flow of technology* (Cambridge, MA: MIT Press).
10. Hillier, *Space is the machine*.
11. M. Weber, *The theory of social and economic organisation*, Eng. edn (Glencoe, IL: Free Press, 1958).
12. H. Spencer, *The principles of sociology* (London: Williams and Norgate, 1877).
13. B. Russell, *The problems of philosophy* (Oxford: Oxford University Press, 1911).
14. B. Hillier, J. Hanson and H. Graham, 'Ideas are in things: An application of the space syntax method to discovering house genotypes', *Environment and Planning B: Planning and Design* 14 (1987): 363–385; Hillier, *Space is the machine*.
15. B. Hillier, A. Penn, J. Hanson, T. Grajewski and J. Xu, 'Natural movement: Or, configuration and attraction in urban pedestrian movement', *Environment and Planning B: Planning and Design* 20 (1993): 29–66.
16. Hillier, *Space is the machine*.
17. B. Hillier, 'Space as a paradigm for describing emergent structure in strongly relational systems', in *Proceedings of the Second International Space Syntax Symposium*, Brasília, edited by F. Holanda, L. Amorim and F. Dufaux (1999), pp. 1–16.
18. B. Hillier and J. Hanson, *The social logic of space* (Cambridge: Cambridge University Press, 1984).
19. A. Giddens, *The constitution of society: Outline of the theory of structuration* (Cambridge: Polity Press, 1984).
20. Respectively, Hillier et al., 'Natural movement'; B. Hillier, 'A theory of the city as object: Or, how spatial laws mediate the social construction of urban space', in *Third International Space Syntax Symposium*, edited by J. Peponis, J. Wineman and S. Bafna (Atlanta, GA: Georgia Institute of Technology, 2001), 2.01–2.28.
21. De Sola Pool and Kochen, 'Contacts and influence'.
22. Lee and Devore, *Man the hunter*.
23. Turner, *Schism and continuity*.
24. Turner, *Schism and continuity*.
25. M. Fortes, *The web of kinship among the Tallensi* (Oxford: Oxford University Press, 1959).
26. Hillier and Hanson, *The social logic of space*.
27. E. R. Wolf, *Peasants* (Englewood Cliffs, NJ: Prentice-Hall, 1966).
28. Hillier, *Space is the machine*.
29. L. Wirth, 'Urbanism as a way of life', *The American Journal of Sociology* 44 (1938): 1–24.
30. M. Sahlins, *Stone Age economics* (London: Tavistock Publications, 1974).
31. Quoted in M. Douglas, *Natural symbols* (London: Barrie & Rockcliffe, 1970).
32. Fischer et al., *Networks and places*.
33. Hillier and Hanson, *The social logic of space*.
34. L. Benevolo, *The origins of modern town planning*, translated [from the Italian] by Judith Landry (London: Routledge and Kegan Paul, 1967).
35. See, on London, M. Young and P. Willmott, *Family and kinship in East London* (Harmondsworth, Middlesex: Penguin, 1962); on Boston, W. H. Michelson, *Man and his urban environment: A sociological approach* (Reading, MA: Addison-Wesley, 1970); and on cases around the world, B. Hillier, M. Greene and J. Desyllas, 'Self-generated neighbourhood consolidation in informal settlements', *Urban Design International* 5 (2000): 61–96.
36. Hillier and Hanson, *The social logic of space*.
37. These issues are dealt with in the postscript to Hillier and Hanson, *The social logic of space*.
38. Allen, *Managing the flow of technology*.
39. B. Hillier and A. Penn, 'Visible colleges: Structure and randomness in the place of discovery', *Science in Context* 4 (1991): 23–49; A. Backhouse and P. Drew, 'The design implications of social interaction in a workplace setting', *Environment and Planning B* 19 (1992): 573–584.
40. M. G. Brown, 'A theory for the spatial design of organizations' (MSc thesis, University College London, 1982).
41. Granovetter, 'The strength of weak ties'.
42. Hanson and Hillier, 'City of London'.

19 Studying cities to learn about minds (2012)

Space syntax and spatial cognition: Some hits and misses. Introduction to 'Studying cities to learn about minds'

Daniel R. Montello

In his 2012 article in *Environment and Planning B*, 'Studying cities to learn about minds: Some possible implications of space syntax for spatial cognition', Bill Hillier adds to the case for space syntax as one of the richest and broadest conceptual and analytic frameworks in the multi-disciplinary study of human–environment interrelationships. Here, cognitive interrelationships are highlighted. Yes, environments help shape minds and minds help shape environments, both mediated in part by perceptual-motor processes. In this article, the environment is the built (anthropogenic) environment, specifically cities, but a solid argument can be made that the reciprocal causal arrow between minds and environments includes natural environments. This is because the environment, as a theoretical entity in human–environment studies, is not a sentient-free collection of patterned physical matter and energy, but patterned matter and energy organised and interpreted by the top-down activity of brains/minds operating on perceptual-motor input. As Hillier states below, 'we routinely interact intuitively with . . . abstract spatial schemes in the real world'. In any case, the article's opening quotation from Peterson and colleagues points to Hillier's interest in one of the central debates about human knowledge that has arisen over the millennia: epistemological innatism versus empiricism. Is mind something we bring to the world and use to impose order on it, or is it derived from our experience in the world? Hillier argues that several properties of city layouts reflect top-down imposition of structure by the human mind, echoing innatism.

Reciprocal causality is reminiscent of another phenomenon with a long history in the study of space syntax: ambiguous causality. When one observes patterns of human behavioural data that correspond to patterns of environmental structure, either the environmental structure causes the behaviour (as space syntax mostly has it), the behaviour causes the environmental structure, each causes one another or some other variable causes each. Does the path structure of a city cause people to travel more to some places than others, do the services located at certain places because of the city's structure attract people more to some places, both, or something else? In this article, Hillier seems to recognise this ambiguity (and I am confident he became aware of it during his career), but I believe he continues to side with the causal arrow of structure causing behaviour (and, in this article, mind): 'So we have brought to light . . . that the grid configuration itself is largely responsible for the pattern of movement flows along streets.'

In this article, and elsewhere, Hillier and other space syntax proponents make some notable claims about the geometry of mental space. For instance: '[T]he way we navigate spatially is guided not by metric distance . . . but by geometrical and topological factors.' Terminologically, he suggests that we define *geometric* environments 'in terms of straight lines and 90° or 45° angles' and *organic* environments 'in terms of the lack of either'. I think his distinction is really about the schematic simplification of spatial patterns, not about whether a spatial pattern is geometric or organic. Geometry is the study of spatial patterns as formal systems, and includes any and all such patterns, many of which do not include straight lines, right angles or perfect diagonals – curves and fractal shapes are part of geometry, for instance. Geometries run the gamut from topological to projective to affine to metric. At the same time, organic processes (not carried out globally by human intention) can result in straight lines and right angles (frozen lakes, crystals, horizons of prairies). Nonetheless, Hillier clearly has a largely valid point: no one who views large circles on the plains from their airplane seat can fail to appreciate the intentional human act of centre-pivot irrigation. I find very intriguing Hillier's conclusion about the cognition of spatially schematised versus non-schematised environments: 'From inside, we often find the organic easier to understand, from outside the geometric.' Especially the first claim calls for research!

The meaning of geometric aside, I have long believed that space syntax undervalues the role of metric space in human spatial cognition.[1] The recent debate about cognitive *maps* versus *graphs* recognises that the existence of metric spatial knowledge does not presuppose great precision nor comprehensive 2D coverage.[2] Either way, the issue of the geometric sophistication of spatial knowledge and reasoning points to an important omission in this article and elsewhere in the space syntax literature. Individual people differ from each other, cognitively, physically, emotionally and elsewise. We can appreciate this not so much as a miss but as a rich opportunity for fruitful space syntax research that mostly has yet to be harvested. Groups or categories of people differ on average, but most fundamentally, individual people differ from each other (of course, we even differ from ourselves at different times in our existence). Differences in spatial behaviour and cognition include everything from body size and motor abilities to hormone levels, working-memory capacity, mental processing speed, perspective-taking ability, tendency to think in terms of routes or configurations (the former being less metric and 2-dimensional than the latter), preference for certain approaches to solving problems, interest in certain kinds of experiences and activities, training at map reading, tendency to use navigation systems, culturally held preconceptions, and more.[3] Surely, all people do not respond cognitively to urban structure in the same way.

Studying cities to learn about minds

Some possible implications of space syntax for spatial cognition

Bill Hillier

Originally published: Hillier, B. 2012. 'Studying cities to learn about minds: Some possible implications of space syntax for spatial cognition'. *Environment and Planning B: Planning and Design*, 39, 12–32.[4]

Some, but not all, of the spatial maps identified by neurobiological and behavioural research impose a structure that goes beyond, and in consequence alters, our interpretation of the information available in the input alone. For example, the hippocampus appears to impose a Euclidean framework onto non-Euclidean inputs.[5] We propose that in 'distorting' the sensory inputs these spatial maps may impose an order and a structure that our spatial conceptual representations require. We point out the importance of a careful analysis of the intrinsic 'organising factors' that interact with spatial information to structure our knowledge of the spatial world. These organising factors act like a kind of 'syntax' in accord with which inputs to spatial systems are ordered, and in doing so they contribute meaning to the spatial representations themselves.[6]

Spatial cognition and the city

What can we learn of the human mind by examining its products? This paper examines the potential of space syntax to throw light on spatial cognition as an instance of the question. At the urban level, space syntax is a theory of description for urban space that is testable by observation of morphological and functional patterns. To the extent that its descriptions correlate with observable regularities in the spatial morphology and functioning of cities, the theory will be successful. However, both the morphological and functional patterns that we find in cities involve the agency of human minds, so the question arises: to what extent can we learn about human minds by examining these patterns? Here, we examine key aspects of space syntax research about the city from a cognitive point of view, and propose two theoretical conclusions. First, that through an examination of the geometric form of the city we can see evidence of the role played by geometric intuitions in its construction. Second, that the relation between human minds and the physical and spatial city is mediated by spatial laws which are intuitively known to people in the same sense that we intuit physical laws in everyday behaviour. It is argued that the city is a prime example of the imposition of a mental geometric order on space, and a general mechanism by which this occurs is proposed.

The ideal and the organic

We begin with some observations about the city and its geometry. The most basic distinction we make about the form of cities is between the ideal and the organic.[7] The ideal are geometric, the organic are not – or seem not to be. The geometric, we define in terms of straight lines and 90° or 45° angles, the organic we define in terms of the lack of either. The ideal seems to be top-down impositions of the human mind, the outcome of reason, often in association with power (Figure 19.1). We easily grasp their patterns when seen 'all at once'. The organic, we take to be the outcome of unplanned, bottom-up processes reflecting the practicalities of everyday living rather than the ordering of human minds. We cannot easily grasp their patterns when we see them 'all at once' from above. But, curiously, when we walk about in them, and so see them a bit at a time, the very differentiation of their parts can make them easier to navigate than patterns in which parts tend to be similar and in similar relations. From inside, we often find the organic easier to understand, from outside the geometric.

In fact, the relations between 'organic' and 'geometrical' are much more complex than any simple typology. Most cities combine the organic and the geometric in different phases of growth. Rome, for example – like many cities, including London – began organically and grew to be more geometric as it expanded. Tokyo, again like many cities, began geometrically and grew to be organic. We also find cities with bewildering juxtapositions of differently orientated and differently shaped grids, a kind of organic mix of geometric grids. Cities often enclose patches of organic settlement as they grow, and of course we also find geometric interventions in organic grids. Cities then seem hugely different in the way in which their grids are put together. It is something of a surprise, therefore, to discover that, in spite of these differences, pretty well all cities share certain common geometric properties. To understand this, we first need to represent urban grids in a consistent way. What we have been looking at so far are the *least line maps* of cities developed by space syntax, which are probably the simplest consistent representations of urban grids.[8]

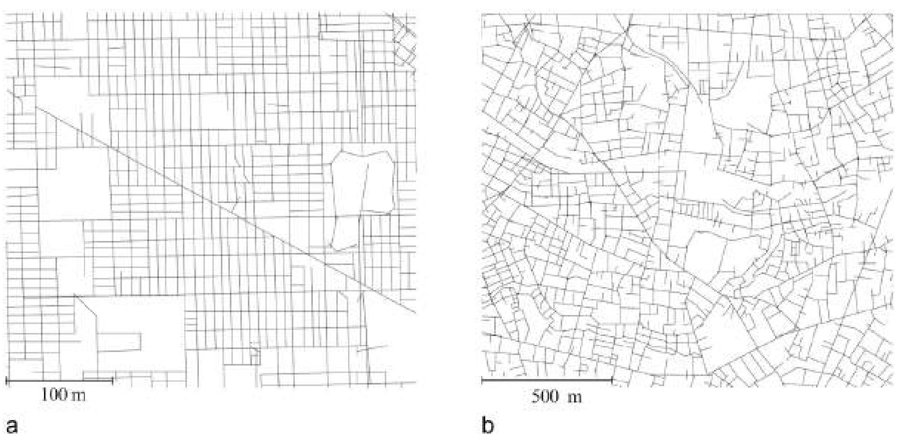


Figure 19.1 Sections of the grids of (a) Chicago and (b) Tokyo.

Some consistencies

With least line maps for real cities, we find some remarkable consistencies. Most strikingly, at all scales, from the local area to the whole city, and for both 'organic' and 'geometric' cities, we find that least line maps are made up of a very small number of long lines and a very large number of short lines, so much so that, in terms of the line length distributions in their least line maps, cities have scale-free properties.[9] Practically speaking, we also find that, wherever we are, we are not far from a line much longer than the one we are on. Formally, this means that these often haphazard urban growths have acquired some mathematical structure. This poses a puzzle. How can mathematically well-formed networks emerge from decades or centuries of activity by innumerable uncoordinated agents acting in very different social, economic and cultural situations, and working with very different, and highly variable, geometries?

But this is not all we find. If we look carefully at organic grids, we begin to see geometrical consistencies. For example, if we look at the section of Tokyo in Figure 19.1, the eye intuitively picks out *line continuities*, formed by series of nearly straight end-of-line connections. If we move along one of these lines, another is very likely to be found at the end of the line, and then another. This happens at all scales, but at each scale the lines are locally longer than lines which lack this kind of angular connection. It can be said that, probabilistically, the longer the line, the more likely it is to end in a nearly straight connection to another line.

The eye will also note a large number of shorter lines, often with near right angle connections, forming more grid-like local patterns. Again, if one such line is found, then it is likely that there will be several others in the immediate neighbourhood. We can also say, the shorter the line, the more likely it is to end or intersect in a right angle or near right angle. These are the opposite properties to those we find in highly formal cities, like Brasília or pre-Columbian Teotihuacan, where the longest lines end at right angles on the most important buildings.

So organic grids tend to have a kind of informal geometry. They are more regular than they appear at first sight. There is, in effect, a hidden geometry in organic cities: they are quite grid-like, in spite of seeming irregular.[10] We can call them *deformed grids*. At the same time, geometric grids are not so regular. Lines are of very different lengths and connectivities, because many are *interrupted*, either by buildings or other artefacts. We can perhaps speak of two kinds of grid: *deformed grids* and *interrupted grids*, as in the sections of Chicago and Tokyo shown in Figure 19.1.[11]

Where, then, do the surprising regularities of both types of grid, and especially organic grids, come from? Three types of process identified in space syntax studies seem relevant. The first refers to how spatial patterns emerge from the simple process of aggregating buildings in space.[12] The second refers to the effects, on the geometry, of the emergent space of how buildings are shaped, scaled and oriented.[13] The third refers to the effects of the emergent spatial geometry on movement patterns.[14] Each process, it is argued in space syntax, is subject to simple spatial laws which can be formally expressed as well as intuitively understood at an everyday level. We can term these the *aggregative* law, the *spatial* law and the *movement* law. We now take each of these in turn.

An aggregative law

The basic form of all cities is one of discrete groups of contiguous buildings, or 'blocks', usually outward facing, defining a network of linear spaces linking the buildings. How can this arise?

If we take cell dyads (Figure 19.2, top left), representing buildings linked by entrances to a bit of open space, and aggregate them randomly, apart from a rule that each dyad joins its bit of open space cell to one already in the system (forbidding vertex joins for the buildings, since no one joins buildings corner to corner), a pattern of buildings and spaces emerges with the topology of a city – outward-facing blocks defining a linking network of linear space – but nothing like its geometry, in spite of being constructed on a regular grid.[15] The 'blocks', and so the spaces, are the wrong shape. Where then does the characteristic urban geometry come from?

A spatial law

This brings us to the second process identified in space syntax. To understand this, we need first to think a little about the network of space in cities and how we interact with it, and the role that different notions of distance might play. Space in cities is about seeing and moving. We interact with space in cities both through our bodies and through our minds. Our bodies interact with the space network by

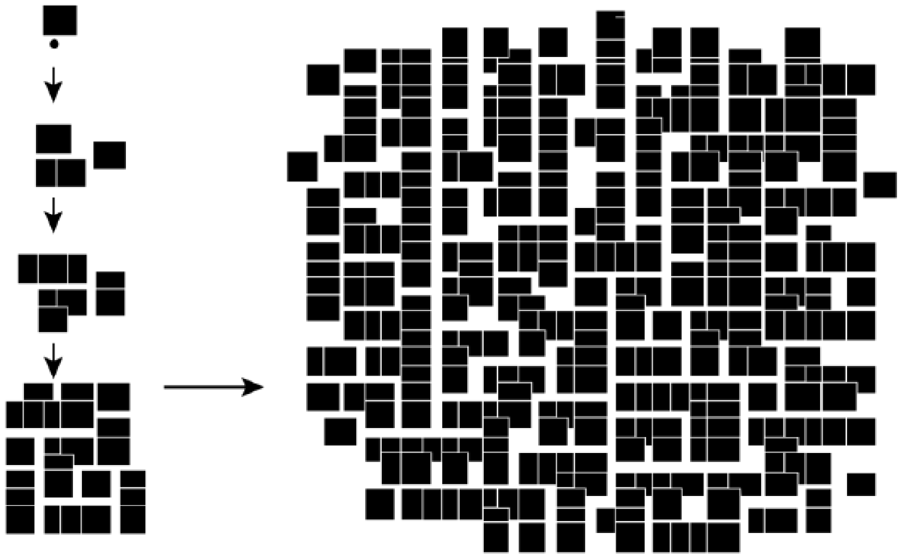


Figure 19.2 The aggregation of dyads of open and closed cells by a restricted random process (vertex joins are forbidden).

moving about in it, and in bodily terms the city exists for us as a system of *metric distances*. Our minds interact with the city through seeing. By seeing the city, we learn to understand it. This is not just a matter of seeing buildings. We also see space, and the city comes to exist for us also as a visually more or less complex object, with more or fewer visual steps required to see all parts from all others, and so as a system of *visual distances*. This warns us that distance in cities might mean more than one thing.

But we also need to reflect on the fact that cities are also collective artefacts which bring together and relate very large collections of people. The critical, spatial properties of cities are not then just the relation of one part to another, but of *all parts to all others*. We need a concept of distance which reflects this. We propose that, if *specific distance* means the common notion of distance as the distance, visual or metric, from *a* to *b* – that is from an origin to a destination – then *universal distance* means the distance from each origin to all possible destinations in the system, and so from all origins to all destinations.[16]

Why does this matter? Because universal distance behaves quite differently from the normal metric and geometric concepts of distance that we use habitually. For example, if, as in Figure 19.3, we have to place a cell to block direct movement between two cells, the closer we place it to one of the outer cells, the less the total distance from each cell to all others will be, because more cell-to-cell trips are direct and do not require deviations around the blocking object.

The same applies to inter-visibility from all points to all others (Figure 19.4). As we move a partition in a line of cells from centre to edge, the total inter-visibility

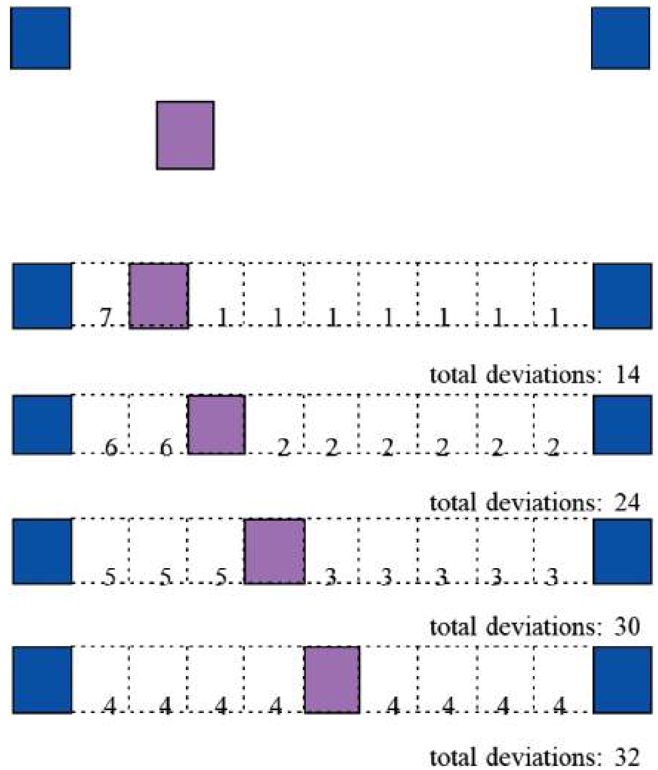


Figure 19.3 Moving an object between two others from edge to centre increases the sum of distances from all cells to all others. The end cells are coloured blue, and the blocking cells are pink. The number in each cell indicates how many deviations from a direct path are necessary to reach all other cells.

from each cell to all others increases, although, of course, the total area remains constant.

Both metric and visual effects arise from the simple fact that to measure inter-visibility, or interaccessibility, we need to square the numbers of points on either side of the blockage. So, all we need to know is that twice the square of a number n will be a smaller number than $(n-1)^2+(n+1)^2$ and that, in general: $2n^2<(n-1)^2+(n+1)^2$.

We can call this the 'squaring law' for space. It applies when, instead of being interested in, say, the distance from *a* to *b*, we are interested in the distance, metric or visual, from each point in the system to all others. In space syntax these 'all to all' properties are called *configurational* to distinguish them from simple relational or geometric properties.

So why does this matter? Because how we place and shape physical objects, such as urban blocks, in space determines the emergent configurational properties

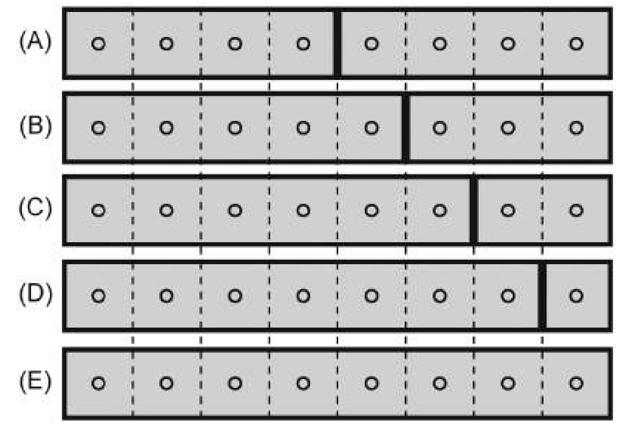


Figure 19.4 Moving a partition from centre to edge increases total inter-visibility. (a) Four points see four cells on both sides, so $2(4)^2=32$ or half of the potential for eight cells; (b) five points see five cells on one side, and three see three on the other, so $5^2+3^2=34$, or 0.53125 of the potential; (c) six see six and two see two, so $6^2+2^2=40$, or 0.625 of the potential; (d) seven see seven and one sees one, so $7^2+1^2=50$, or 0.71825 of the potential; (e) eight see eight, so $8^2+0^2=64$ or all of the potential.

of that space. For example, one consequence of the squaring law is that, as we move objects from corner to edge and then to central locations in bounded space, total inter-visibility in the system decreases, as does visual integration (or universal visual distance) – defined as how few visual steps we need to link all points to all others (Figure 19.5a). The same applies to metric integration (or metric universal distance), defined as the sum of shortest paths between all pairs of points in the ambient space, which increases as we move the obstacle from corner to centre (Figure 19.5b).

The same squaring law governs the effect of shape (Figure 19.6): the more we elongate shapes, keeping the area constant, the more we decrease inter-visibility and increase trip length in the ambient space. The effect of a long and short boundary is to create greater blockage in the system through the squaring law.

Even at this stage, this spatial law has a critical implication for cities: in terms of configurational metrics, a short line and a long line are, other things being equal, metrically and visually more efficient in linking the system together than two lines of equal length (Figure 19.7), as would be a large space and a small space, compared with two equal spaces.

Another consequence is for the mean length of trip (or metric integration) from all points to all others in different types of grid, holding ground coverage of blocks, and therefore total travelable distance in the space, constant. In the four grids in Figure 19.8, cooler colour shading means shorter mean trip length to all

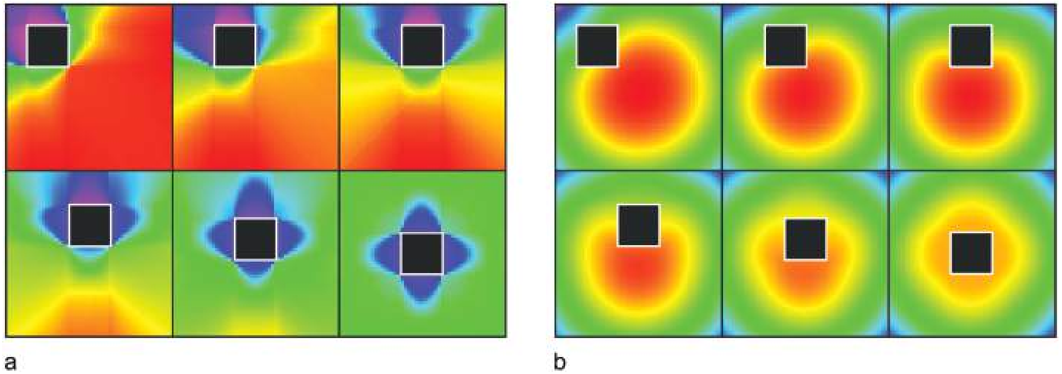


Figure 19.5 (a) Moving an object from corner to centre decreases inter-visibility. Red means less visual distance to all other points, through the spectrum to blue. (b) Moving an object from corner to centre increases the mean length of trips (blue is less metric distance, through the spectrum to red).

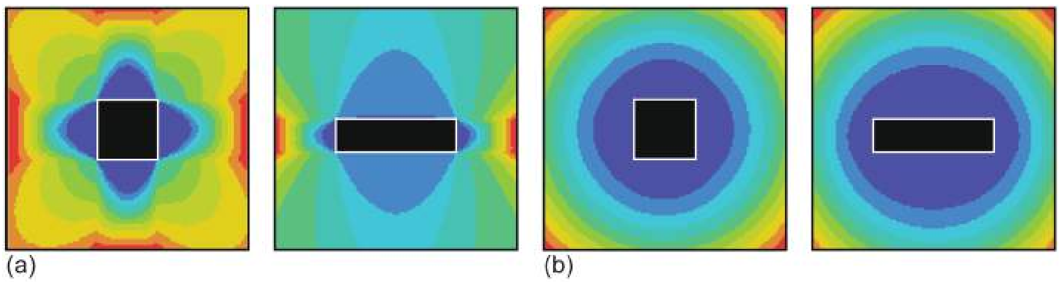


Figure 19.6 Changing the shape of an object from square to rectangular decreases inter-visibility and increases mean trip length. Colours near the red end of the spectrum from red to blue mean (a) less visual distance and (b) less metric distance.

other points. Compared with the regular orthogonal grid (top left), interference in linearity on the right slightly increases mean trip length. But more strikingly, if we reduce the size of central blocks and compensate by increasing the size of peripheral blocks, we reduce mean trip length compared with the regular grid. This, of course, is the 'grid intensification' that we often note in looking at centres and subcentres in cities. As so often, we find a mathematical principle underlying an empirical phenomenon.

How we place and shape objects in space then determines the emergent configurational properties of that space. But what kind of block placing, and shaping, make space urban?

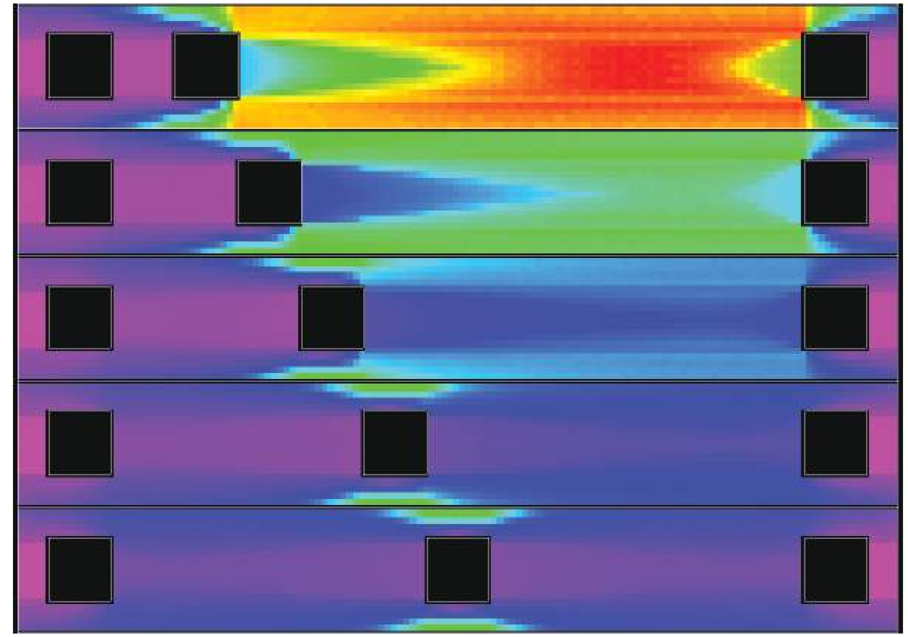


Figure 19.7 Other things being equal, a short and long line integrate more than two lines of equal length. Colours near the purple end of the spectrum from red to purple indicate less visual distance.

In Figure 19.9a buildings are aggregated in an approximately urban way, with linear relations between spaces, so we can see where we are going as well as where we are. In Figure 19.9b, the identical blocks are retained but are moved slightly to break linear connections between the spaces. If we then analyse metric and visual distances within the two complexes, we find that all-to-all metric distances (not shown) increase in the latter case, so trips are, on average, longer, but the effect is slight compared with the effect on all-to-all visual distances, which change dramatically (shown in Figure 19.10). Showing visual integration – darker shading means less visual distance as before – we see that the case in Figure 19.10a identifies a kind of main street with side and back streets, so an urban type of structure has emerged. But the case in Figure 19.10b has lost both structure and degree of inter-visibility. Even though the changes are minor, it feels like a labyrinth. We can see where we are but not where we might be.

The effect on computer agents moving around the system is striking, if obvious. In Figure 19.11, 10,000 computer agents with forward vision are moved in the space, again using the software by Turner.[17] The agents randomly select a target within their field of vision, move three pixels in that direction, then stop and repeat the process. In Figure 19.11a, the traces of agent movement 'find' the

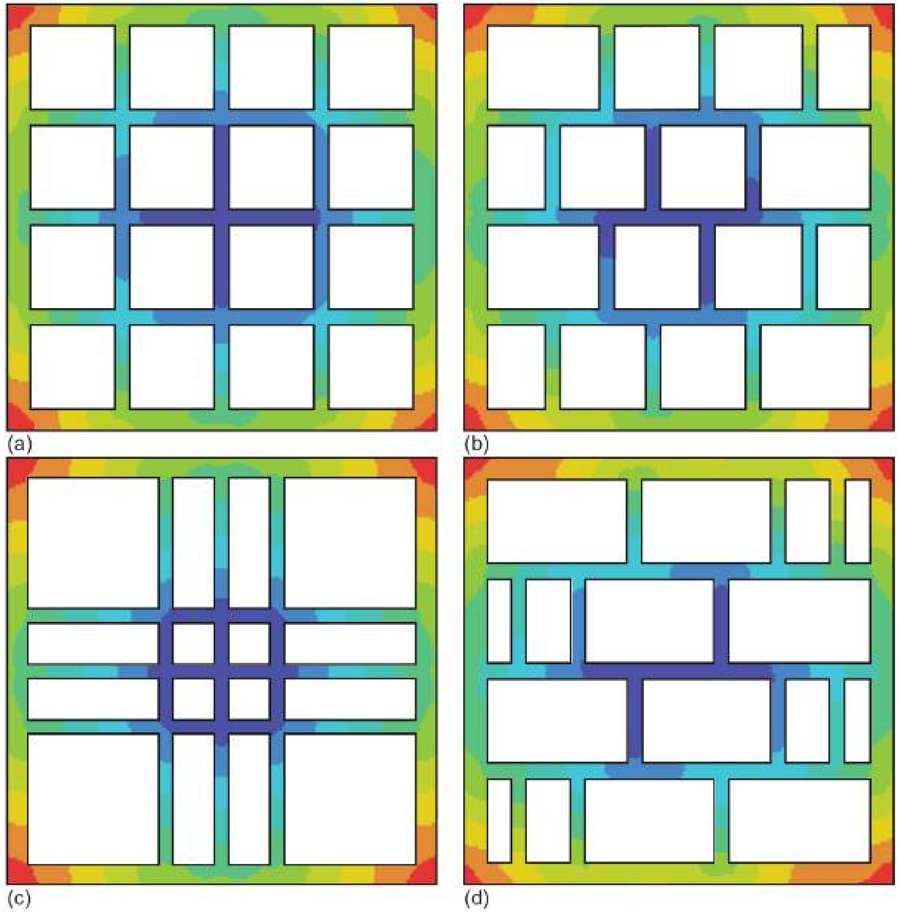


Figure 19.8 Changing the scaling of a grid changes mean trip length. In this case, for graphical clarity, colours near the blue end of the spectrum from red to blue mean less metric distance from each point to all others. The mean distances for each system are: (a) 2.53; (b) 2.59; (c) 2.42; (d) 2.71.

structure of visual integration. In Figure 19.11b they wander everywhere and tend to get trapped in 'fatter' spaces. This is an effect purely of the configuration, since everything else is identical.

But what about human beings? Human beings do not of course move randomly, but purposefully, and successful navigation in an unfamiliar environment would seem to depend on how good a picture of the whole pattern we can get from seeing it from a succession of points within it. One way to plausibly measure this property is by correlating the size of the visual field we can see from each point with the visual integration value (its visual distance from all other

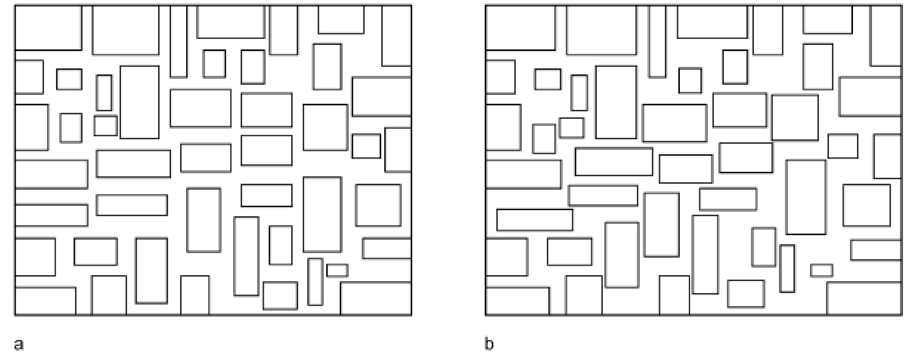


Figure 19.9 Two slightly different arrangements of identical blocks, with (a) strong linear relations between spaces and (b) weak linear relations between spaces.

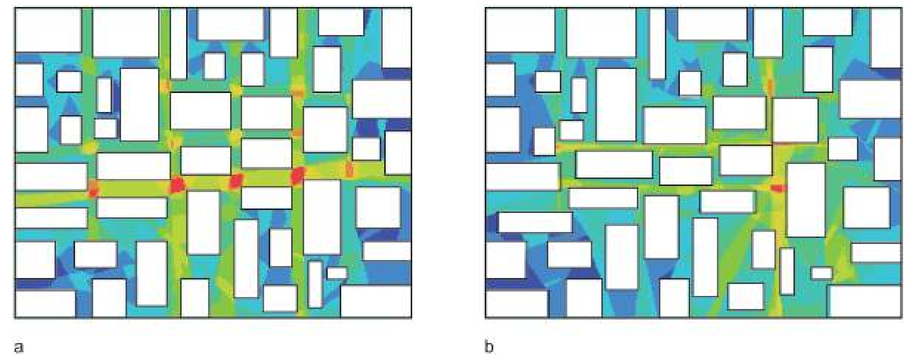


Figure 19.10 Visual integration analysis (colours near the red end of the spectrum from red to purple indicate higher integration, and so low visual distances from all points to all others) showing how a non-urban layout loses both integration and structure through slight block changes.

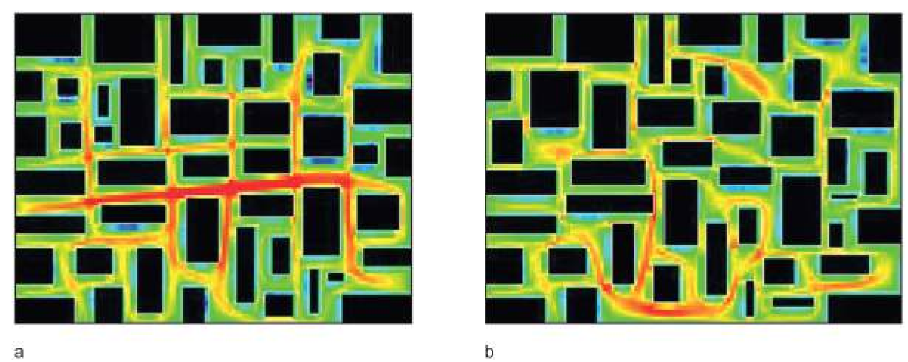


Figure 19.11 Traces of 10,000 forward-looking agents moving nearly randomly in two slightly different arrangements of blocks. Red means many traces, blue few.

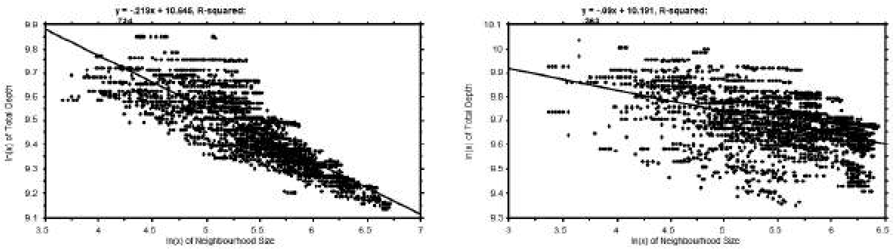


Figure 19.12 Intelligibility scattergrams for the two layouts in Figures 19.9a and 19.9b, respectively.

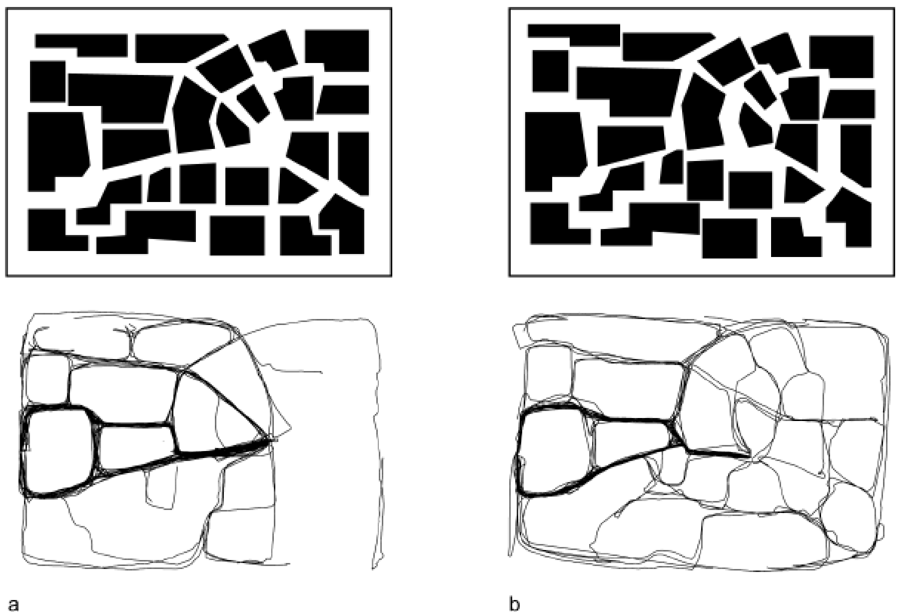


Figure 19.13 Trace of human agents in (a) intelligible and (b) unintelligible layouts. Credit Ruth Conroy Dalton.

points), in effect measuring the relation between a *local* property that we can see from each point, and a *non-local* one that we cannot see (Figure 19.12).

In space syntax this is called the *intelligibility* of the system. The r^2 for the 'intelligible' layout in Figure 19.12a is 0.714 while in Figure 19.12b it is 0.267. Defined this way, the intelligibility of a spatial network depends almost entirely on its linear structure. Field studies and experiments suggest that this does work for humans.[18] For example, Conroy Dalton took a linearised 'urban' type network (Figure 19.13a) and asked subjects to navigate in a 3D immersive world from the left edge to the 'town square' and back again. As the traces show, they managed

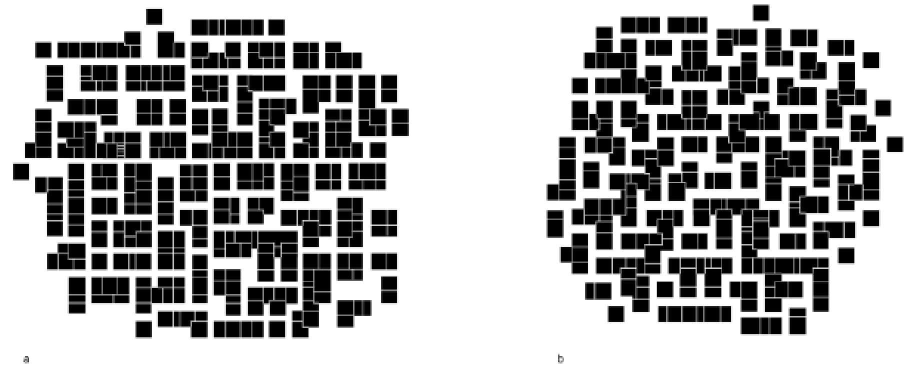


Figure 19.14 (a) A layout generated by a 'conserve longer lines' rule; (b) a layout generated by the inverse rule.

to find reasonable routes. But the (identical) blocks were then moved slightly to break the linear structure and to reduce intelligibility (Figure 19.13b), and the experiment was repeated. The subjects found the modified layout labyrinthine, and many wandered all over the system trying to perform the same wayfinding task.

So, if, coming back to our aggregative process, we modify it by requiring those adding cells to the system to avoid blocking a longer local line if they can block a shorter one (Figure 19.14a), we find a much more urban type of layout emerges, approximating the mix of long and short lines we find in real systems and emulating certain structural features.[19] With the contrary rule – always block long lines (Figure 19.14b) – we construct a labyrinth in which lines are of much more even length. So urban space networks seem to be shaped in some degree by a combination of spatial laws and human agency, with the human agent implementing, and so in a sense knowing, the spatial laws. The consistency we find in urban space patterns suggests that human beings 'know' the configurational laws of space in some sense – perhaps in the same sense that they 'know' simple 'intuitive physics' when they throw a ball of paper so that its parabola leads it to land in a waste paper basket.

A movement law

But this cannot be all. Cities are also shaped by economic and social processes. How do these fit into the picture? This is the third thing we need to know. First, we can note that the 'urban' type pattern in Figure 19.14a is *dual* in the sense that it is composed of a dominant pattern of long lines against a background of areas made up of short lines, approximating what we saw cities to be like. But why this duality? For this we must understand the functional effects of the network, and to understand these we must first learn to analyse the network in terms of what the network is primarily for: that is, movement.

Let us first reflect for a moment on human movement. Spatially speaking, every human trip is made up of two elements:

(1) an origin-destination pair – every trip is from an origin space to a destination space – we can call this the *to-movement* component;
(2) the spaces passed through on the way from origin to destination – we can call this the *through-movement* component.

In fact, both of these potentials can be measured:

- to-movement is about the *closeness* or *accessibility* of spaces from all others – in space syntax this is normalised and called the measure of *integration*: how close a space is to all other spaces;[20]
- through-movement is about the propensity of spaces to be passed through on the way from all origins to all destinations – this is the betweenness measure, called *choice* in space syntax, meaning how likely a space is to be chosen as part of a route between spaces.[21]

Starting from the least line map, each line is divided into its segments (between intersections) and the result is represented as a graph (Figure 19.15). Integration (closeness in mathematical parlance) and choice (betweenness in mathematical) measures are then assigned to relations between each segment and all others, using shortest path (metric), least angle change (geometric) and fewest-turns (topological) weightings, and are applied at different radii from each segment – with radii also defined metrically, geometrically and topologically. This yields a matrix of configurational measures which we can use to see if we can find significant structure-function relations.

So, we can look at each segment in a system in terms of either its to-movement or through-movement potential, defining distance and radius metrically (shortest paths), geometrically (least angular change paths) or topologically (fewest-turns paths).[22]

Taking the least line map of the system generated in Figure 19.14a, we can visualise the pattern of values, which we call the *structure* of the system, by 'colouring up' the network in the usual way for integration (Figure 19.16a) and choice (Figure 19.16b) without radius restriction. Since integration measures the accessibility of nodes as destinations from origins, then, from the principle of distance decay (and other things being equal), we must statistically expect more movement potential for nodes that are closer to all others, at some radius. Likewise, since choice measures the sequence of segments we pass through, we must expect a similar bias in real movement. In effect, integration measures the to-movement potential and choice the through-movement potential of spaces.

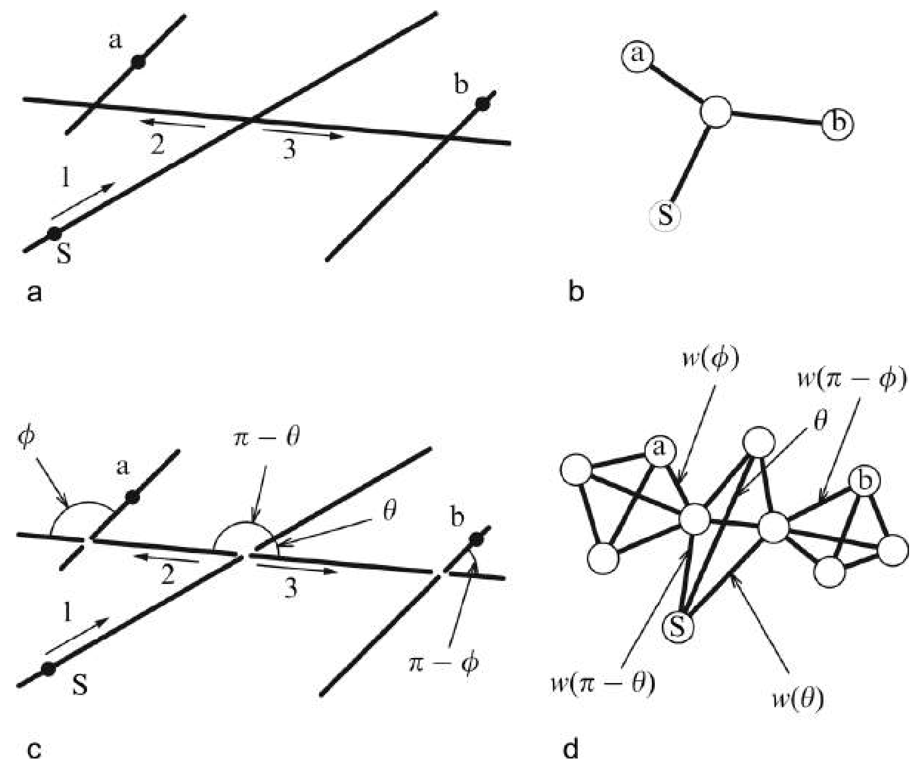


Figure 19.15 How the angular segment map is derived from the least line map: (a) line model; (b) graph of line model; (c) segment model; (d) graph of segment model.

Natural movement

Is there a link to real movement? This will depend, among other things, on the question, much discussed in the cognitive literature, of how people make distance judgements in complex space.[23] So how do they? Shortest paths? Fewest turns? Least angle change? Hillier and Iida suggest this can be resolved by examining real flows.[24] They applied the three weightings to the two measures to make six different analyses of the same urban system, and correlated the resulting patterns of values for each segment with observed movement flows on that segment (Table 19.1), arguing that if, across cases, there were consistently better correlations with one or other weighting, then the only logical explanation would be that this weighting reflects better how people are biasing spatial movement choices, since everything else about the system is identical. In fact, across four separate studies in areas of central London, we consistently found that geometric, or least-angle, weighting

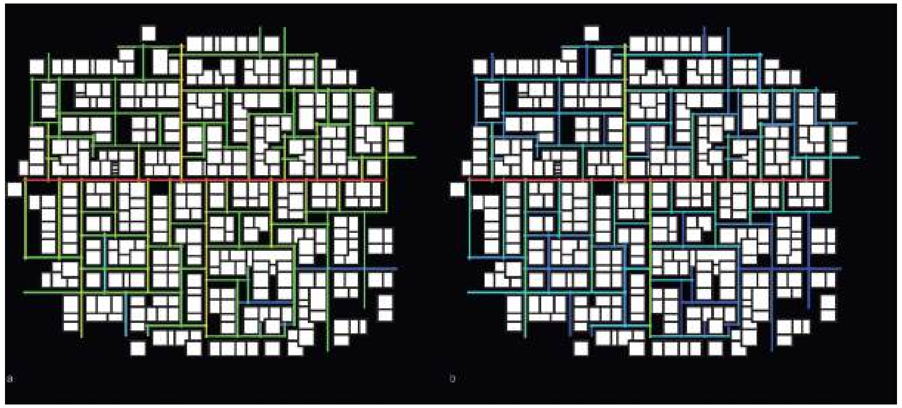


Figure 19.16 Integration, or closeness (a), and choice, or betweenness (b), analysis of the least line map of the system in Figure 19.14a.

yields the strongest movement prediction, with an average r^2 of around 0.7 for vehicular movement and 0.6 for pedestrian, closely followed by the topological or fewest-turns weighting. Metric shortest paths are markedly inferior in most cases, and, in general, to-movement potentials are slightly stronger than through-movement potentials, though this varies from case to case.[25]

This shows that the configuration of the network is responsible for a substantial part of movement flows in two senses: the objective to-movement and through-movement potentials of the network itself contribute what we might call *network effects* on shaping flows; and these are modified by how human minds contribute distance effects through how they read distance in complex spaces. So, we have brought to light two rather remarkable things. The first is that the grid configuration itself is largely responsible for the pattern of movement flows along streets. We call this the theory of natural movement. Second, the way we navigate spatially is guided not by metric distance, as has been uncritically assumed, but by geometrical and topological factors.

The spatial form of cities

With this knowledge, then, we have a new tool for investigating the form and functioning of cities. First let us look at the emergent spatial form of cities. Applying the *integration*, or to-movement, measure to real cities, and using the least angle change, or geometrical, definition of distance, we find some remarkable emergent geometrical patterns, and again we find that they are near invariant across different kinds of city. For example, we commonly find a pattern we call a *deformed wheel*: a hub, spokes and rim forming the main structure of public space, and the residential areas in the interstices of the wheel. This first came to light in the study

Table 19.1 Correlations between respective flows and shortest-path, least-angle and fewest-turns analyses, applied to accessibility and choice measures across four areas of central London

a. Vehicular Movement: r^2 values for correlations between vehicular flows and shortest-path, least-angle and fewest-turns analyses, applied to accessibility and choice measures. Best correlations are marked *. Numbers in brackets indicate best radius in segments for accessibility measures.

	Gates	*Measure*	*Least length*	*Least angle*	*Fewest turns*
Barnsbury	116	accessibility	0.131 (60)	0.678 (90)	0.698 (12)
		choice	0.579	0.720*	0.558
Calthorpe	63	accessibility	0.950 (93)	0.837* (90)	0.819 (69)
		choice	0.585	0.773	0.695
South Kensington	87	accessibility	0.175 (93)	0.688 (24)	0.741* (27)
		choice	0.645	0.629	0.649
Brompton	90	accessibility	0.084 (81)	0.692* (33)	0.642 (27)
		choice	0.475	0.651*	0.588

b. Pedestrian Movement: r^2 values for correlations between pedestrian flows and shortest-path, least-angle and fewest-turns analyses, applied to accessibility and choice measures. Best correlations are marked *. Numbers in brackets indicate best radius in segments for accessibility measures.

	Gates	*Measure*	*Least length*	*Least angle*	*Fewest turns*
Barnsbury	117	accessibility	0.119 (57)	0.719* (18)	0.701 (12)
		choice	0.578	0.705	0.566
Calthorpe	63	accessibility	0.061 (102)	0.637 (39)	0.624* (36)
		choice	0.430	0.544	0.353
South Kensington	87	accessibility	0.152 (87)	0.523* (21)	0.502 (15)
		choice	0.314	0.457	0.526*
Brompton	90	accessibility	0.111 (81)	0.623* (63)	0.578 (63)
		choice	0.455	0.513*	0.516

of small towns in the south of France, and the same pattern was found in London's urban areas with its 'urban villages' at the hub. But it was something of a surprise to find the pattern approximated in very large cities such as London, with a relatively weak rim, and Tokyo, with much stronger and multiple rims (Figure 19.17).

We seem to find this pattern emerging under very different geometric conditions. For example, we find this emergent structure in strongly geometrical Atlanta, and the very ungeometrical, old city of Hamedan in Iran (Figure 19.18). When we apply the *choice*, or through-movement, measure, we find a different kind of structure, reflecting some of the deformed wheel but more like a network

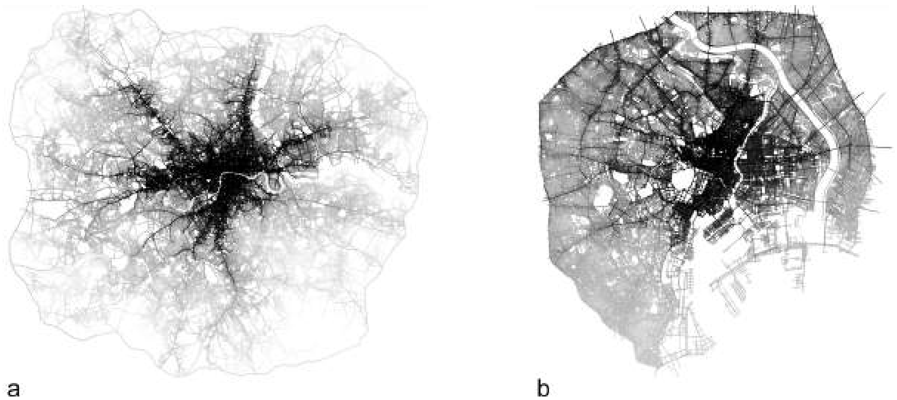


Figure 19.17 Integration analysis of (a) London within the M25 and (b) metropolitan Tokyo, showing the underlying approximation of the 'deformed wheel', in the Tokyo case with multiple rims.

Figure 19.18 Integration analysis of (a) Atlanta and (b) Hamedan showing how the deformed wheel pattern emerges under very different geometrical conditions.

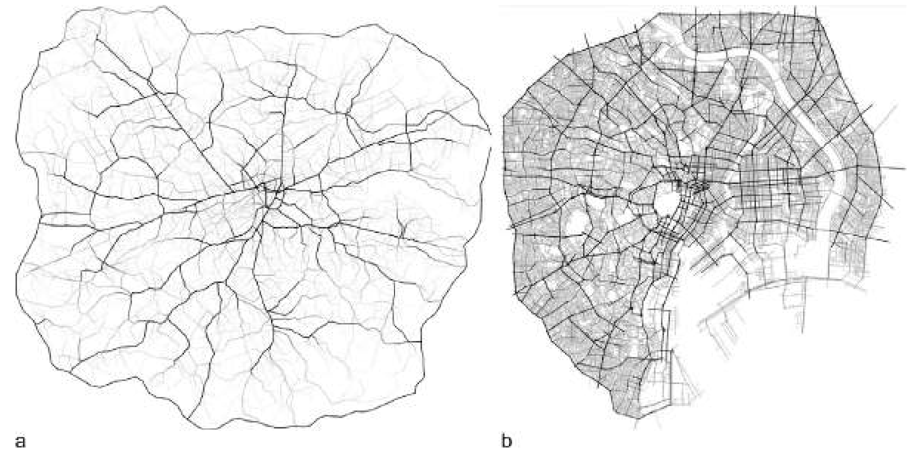


Figure 19.19 Least angle choice (betweenness) analysis of (a) London and (b) Tokyo, showing the network pattern in both cases.

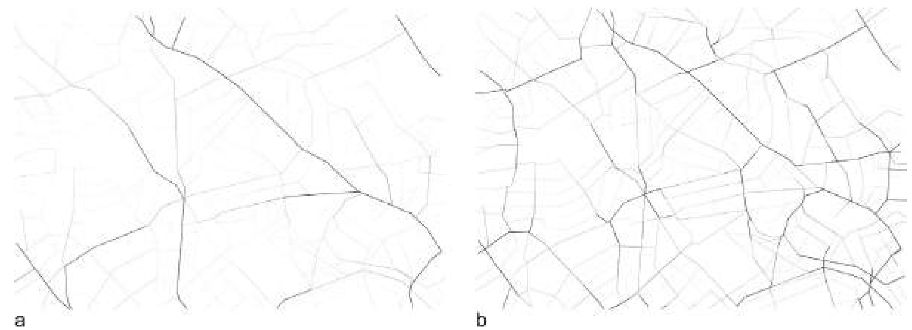


Figure 19.20 (a) Unrestricted radius and (b) radius 1,000 m choice analysis of a part of northwest London.

(Figure 19.19), and again, this seems to work for interrupted grids as well as for deformed grids. We can also combine the two measures by simply multiplying one by the other, to give a combined picture of the to-movement and through-movement potentials of each street segment in the system with respect to all others.

Applying the restricted radius measures then allows us to capture much more detail of local structure, reflecting the fact that when we make large-scale trips in the city we tend to use the global structure, but when we move locally we will often find ourselves prioritising spaces which are not part of a global pattern, but which are locally important. We can use the greyscale range as a kind of microscope to explore these detailed local patterns. For example, in Figure 19.20 we reduce the radius of the choice measure from (a) unrestricted to (b) 1,000 m in an area of northwest London, and then use the greyscale range to zoom in and begin to detect the urban villages, which of course are focuses of local, but not global, movement.

We can use the same technique to detect London's often surprising pattern of local shopping street and market areas, and to identify local area structures, often in the form of a local deformed wheel. The local deformed wheel is, in fact, the secret of London's surprisingly strong local organisation and the reason we name it as a system of 'urban villages'.

But how do these patterns affect the functioning of the city? We already know that the emergent structure of the grid reflects and shapes movement flows. Does this have further consequences? It does, and by understanding these consequences we arrive at a new theoretical definition of the city. What we find is that the link between the network configuration and movement flows is the key to the dynamics and evolution of the system. Because the network shapes movement, it also, over time, shapes land-use patterns, in that movement-seeking land uses, such as retail, migrate to locations which the network has made movement-rich while others, such as residence, tend to stay at movement-poor locations. This creates multiplier and feedback effects through which the city acquires its universal dual form as a foreground network of linked centres and subcentres at all scales into a background network of residential space. This is the space syntax definition of a city. Through its impact on movement, the network has set in train a self-organising process by which collections of buildings become living cities.

The dual city

We have then found our dual structure, and we can explain it. The foreground structure, the network of linked centres, has emerged to maximise grid-induced movement, driven by micro-economic activity. Micro-economic activity takes a universal spatial form and this type of foreground pattern is a near-universal form in self-organised cities. The residential, background network is configured to restrain and structure movement in the image of a particular culture, and so tends to be culturally idiosyncratic, often expressed through a different geometry which makes the city as a whole look spatially different. We call the first the *generative* use of space, since it aims to generate co-presence and make new things happen, and the second we call *conservative*, since it aims to use space to reinforce existing features of society. In effect, the dual structure has arisen through different effects of the same laws governing the emergence of grid structure and its functional effects. In the foreground, space is more random, in the background it is more rule-governed, and so has more conceptual intervention.

We can illustrate this most clearly in a city with more than one culture (now unfortunately separated): Nicosia. In Figure 19.21, top right is the Turkish quarter, bottom left the Greek quarter. Their line geometry is different. In the Turkish quarter, lines are shorter, their angles of incidence have a different range, and there is much less tendency for lines to pass through each other. Syntactically, the Turkish area is much less integrated than the Greek area. We can also show that

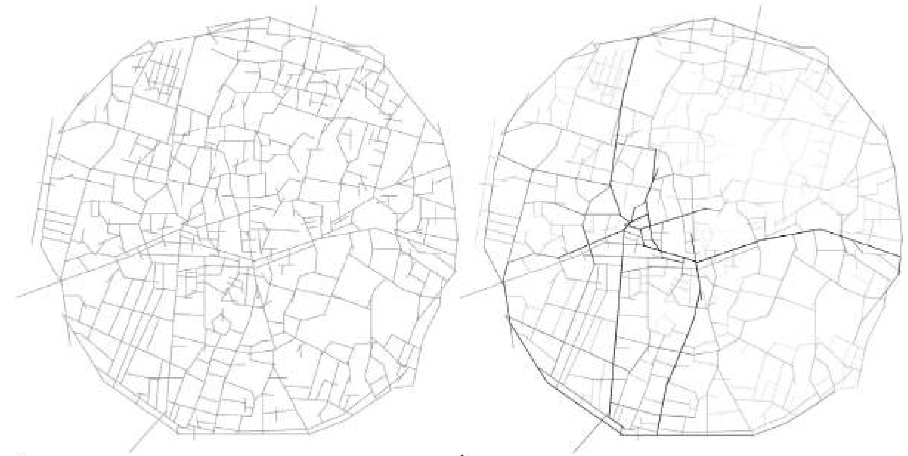


Figure 19.21 (a) The old city of Nicosia and (b) its integration analysis, showing the deformed wheel core in spite of culturally differentiated residential space.

it is less intelligible, and has less synergy between the local and global aspects of space. Yet, in spite of these strong cultural differences in the tissue of space, we still find Nicosia, as a whole, is held together by a clear deformed wheel structure. This shows how micro-economic activity spatialises itself in a universal way to maximise movement and co-presence, while residence tends to reflect the spatial dimension of a particular culture, and the expression is in the first instance geometrical. Since residence is most of what cities are, this 'cultural geometry' tends to dominate our spatial impressions of cities.

A cognitive conjecture: How do we acquire non-local knowledge of the city?

But whatever their geometry, space networks in cities have a further unexpected property. Although the form of the system has evolved bottom-up, its functioning is top-down, in that the movement flows which drive the evolution of the system reflect the position of each space in the large-scale configuration, not the local properties of the space. In this sense, the properties of spaces which are critical to its functioning are *non-local* and reflect a large number of remote connections. This poses a very interesting cognitive question. In order to produce the patterns of flows we find, people must be using some kind of non-local internal representation of the space network, with both geometrical and topological properties. Since people see cities only a bit at a time, it seems that they somehow *synchronise* discrete experiences into a *non-local* picture. What might this be like?

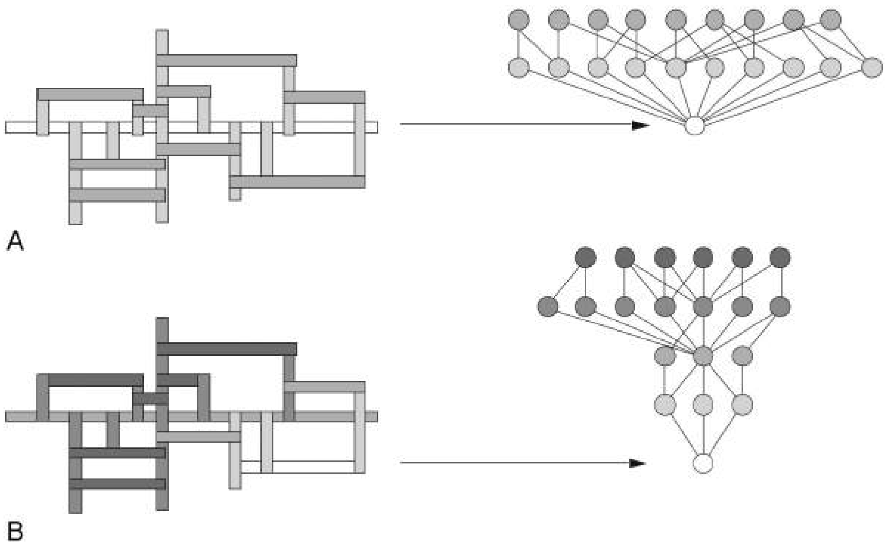


Figure 19.22 A notional grid and two of its justified graphs: (a) total depth from root = 28; (b) total depth from root = 54.

In space syntax, non-local patterns are visualised through a graphical device we call the *justified graph*, or *j-graph*.[26] We illustrate this through Figure 19.22. Each element in the system – in the case in Figure 19.22 each 'street' – is represented as the root of a graph with all other elements to which it connects aligned in the first layer above, those that connect to this layer in a second layer, and so on. In this way we can picture, for example, the complexity of routes from each street to all others, a critical non-local property. A graph shallow from the root, called an *integrated* j-graph in space syntax, means little route complexity to all other lines, a graph deep from the root indicates much complexity.

The representation that people are using for navigation looks as though it might be some kind of justified map of connected lines: a *justified line graph*, perhaps. What might cognitive science say about this? First, it says that human beings have *egocentric* route knowledge and *allocentric* map knowledge, and that these are actually dealt with in different parts of the brain.[27] It is the latter we need to solve navigational problems beyond the scope of simple route recall. But cognitive science also says something else very interesting. In memory, people routinely correct bends to lines, and near right angles to right angles.[28] These are instances of what we might call the 'Kantian simplification': people impose more geometry on the situation than it actually has.[29] But in terms of our putative line graph, correcting lines is not so much a geometrical error as a topological simplification. Its effect is to *turn two line elements into one* and so simplify the line topology. It also eliminates an asymmetric relation by which we must go through this space to get to that one, so making the j-graph shallower.

We might then conjecture that going from route knowledge to map knowledge is a matter of going *from the j-graph* (in fact its spanning tree) *to the graph*. The 'Kantian simplification' may be part of the means by which we pass from an egocentric to an allocentric understanding of space, since it makes the topogeometric j-graph simpler and less asymmetric, and so less complex in two critical ways. The more we straighten lines and correct turns, the more the graph becomes shallow from its root, symmetric and bipartite. This makes the line graph easier to transform to see other viewpoints, and it is perhaps this transformability that permits the passage from an egocentric to an allocentric model. Could this be how we learn cities?

Further cognitive reflections on the objective city

Two further aspects of this account of cities stand out as raising cognitive challenges. Both reflect the fact that the city, and its functioning, as we have described them, are products of human agency, and inconceivable without the active role of human cognition. The first is that the complex emergent patterns and processes we have described as near-invariants in the form and functioning of cities are constructed from the metric and geometric properties of the system of linear spaces that links the city into a single system in the first place – that is, the lengths of lines and the angles of incidence at the intersections between lines. The line graphs which have been the basis of our analysis are composed, essentially, of no more than this.

We find, in effect, that the city has a pervasive two-level geometry. There is a geometrical consistency in the way in which the relations between line lengths and angles of incidence at the intersection are formed into local patterns to give the local differentiation between the more public and more residential parts of the system. And there is a geometrical consistency in the way in which these patterns are scaled up to give the emergent structure of the whole city. It is these geometrical consistencies that govern the form-function relation in cities, and which relate the spatial patterning of the city with the emergent spatial patterning of city life. The city cannot have acquired the pervasive *geometricity* of its form without the equally pervasive involvement of human geometric intuition. In this sense, human geometric intuitions seem to be *embedded* in the city itself.

Second, we have also described a system in which *spatial laws*, implemented through human agency, intervene between the patterns of micro-economic and social activity that animate the city, and the emergent spatial form of the city itself. We have no concept for such a *human-mediated* but *law-governed* system. But what is clear is that the relation between human activity and space is mediated by spatial laws, and it is only through the lawfulness of space that economic and social activity is able to express itself in space. On reflection, it must be that way.

An unavoidable consequence of this is that cities are, in a profound sense, *cognitive* – and so human-objects before they are economic and social objects.

Paradoxically, it is because they are so that they can absorb, without strain, massive changes in the patterns of activity that animate them. The relation between the form of the city and its functioning is, in effect, *generic*, not specific. It is not specific patterns of activity that shape space but the way in which the relation between space and activity is mediated by the need for movement and co-presence. New patterns of activity, like the old, will require the full continuum of spaces, from integrated to segregated, and so will discover how to fit into and perhaps modify slightly an existing urban pattern. This *capaciousness* of the city for the absorption of new functional patterns comes from the underlying form of the city – that is, the *generic cognitive* city – and its priority over the more specific social and economic city.

A generic mechanism: Description retrieval?

The *signs of minds* that we detect in the city, then, suggest the pervasive involvement both of geometric intuitions and of spatial laws, both in the formation of the city and in its functioning. Is there a general mechanism governing this link? Here it is proposed that there is, and that it depends on the proposition that our mental interaction with the spatial world engages abstract relational ideas as well as concrete elements. In general, spatial relations are ideas we think *with* rather than *of*. The classic case is the *prepositions*, like *between* or *beyond*, all of which embody bundles of relations in abstract form, and we use them routinely to structure our picture of the world. It is proposed that there is a generic mechanism, called *description* retrieval,[30] through which we extract abstract information from concrete events and re-embody it in real time.

Suppose, for example, as in Figure 19.23, one person (a) builds a house and another person (b) builds a house next to it. In one case, subsequent builders ignore the first relation [(c) and (d)] and create a random pattern; in another [(e) and (f)] other builders 'get the idea' of the initial relation and re-embody it in subsequent actions, and a regular form emerges.

This process is interesting at two levels: that of the process and that of the emergent form. The relation 'next to' has an interesting abstract property: it is *symmetrical*, in that, if *a* is next to *b*, then *b* is next to *a* – unlike, for example, *above* or *below*, or *behind* and *in front*, which are *asymmetrical*. In reproducing the relation, then, builders are reproducing the abstraction, and in this sense emulating a rule-following behaviour. We can think of this as a kind of embedded *rule*: the rule is inherent in the concrete behaviour. Human behaviour is full of embedded rules of this kind, and often what is embedded is an abstract scheme of spatial relations. It is hard to avoid the inference that we routinely interact intuitively with the abstract spatial schemes in the real world. The emergent forms [Figures 19.23(d) and 19.23(f)] are also interesting. In the regular case, where the rule following has put similar things in similar relations, we easily retrieve a description of the high-level object as a line or perhaps 'terrace', but in the random case we do not. In the

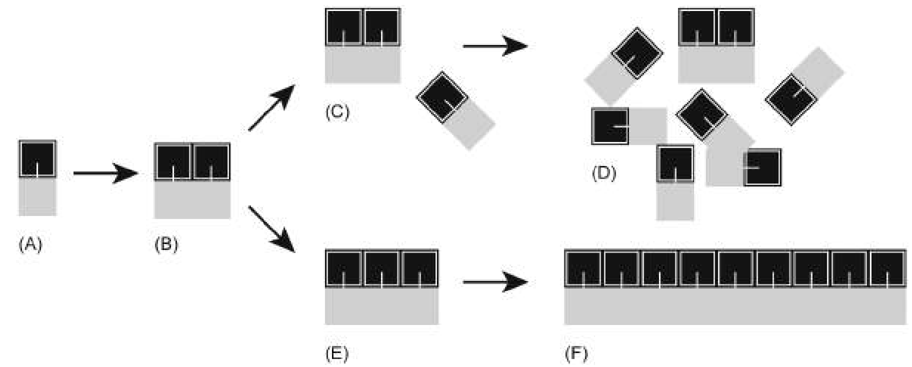


Figure 19.23 The generation of simple forms from simple rules. See text for a description of (a)–(f).

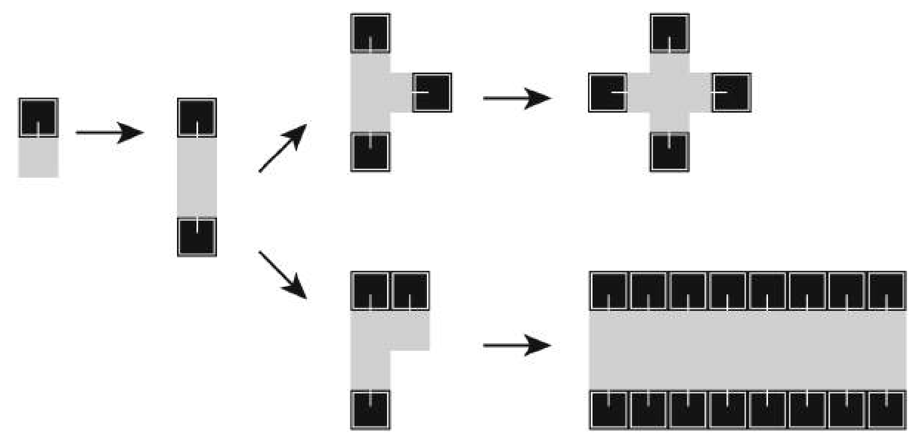


Figure 19.24 The generation of other simple forms from the use of two embedded rules.

regular case, it seems we are able to *synchronise* the discrete objects into an overall shape to create a higher-level *template* for further embodiment. In the other case, we cannot. In general, this seems to be the case.

By following two embedded rules, rather than one (Figure 19.24), we find other cases in which local rule following leads to the emergence of other regular global forms – one we would name a *courtyard*, the other a *street*. In fact, the basic forms for which we have names can all be seen as products of a local rule following, leading to global forms in which *similar things are put into similar relations*.

But exactly what is it that we recognise at this *template* level? Is it just the *recursion process* of locally similar events? Or is the *upper-level* recognition, in some

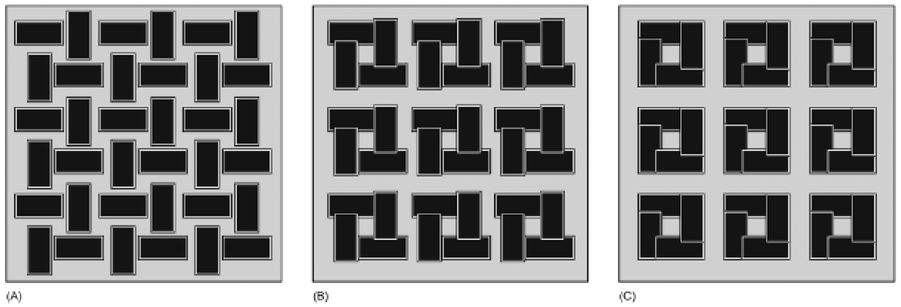


Figure 19.25 By increasing the linear organisation from left to right we move from retrieving a local to a global description.

sense, independent of the lower-level constructive process? The cases in Figure 19.25, in which the same objects are rearranged to allow more global linear connections, provide the answer: the upper level is *independent* of the lower level. Figure 19.25a is no less recursive than Figure 19.25c, but we see it only as a recursive local pattern, not as a whole. Figure 19.25b we are more or less able to synchronise as a whole because lines have appeared at the level of the whole. In the case of Figure 19.25c, we are far more aware of the global pattern than of any recursive process. We have a kind of *global takeover* and we retrieve an upper-level *template*.

So, the process of *synchronisation* in the upper level of the *description retrieval* process seems to require certain formal properties to be satisfied at that level, and will not happen without those properties. We begin to see why the emergent spatial complexity of the city *requires* a certain *emergent geometry* if human beings are to interact with it effectively.

Notes

1. An argument I develop in D. R. Montello, 'The contribution of space syntax to a comprehensive theory of environmental psychology', in *Sixth International Space Syntax Symposium*, edited by A. S. Kubat, Ö. Ertekin, Y. I. Güney and E. Eyüboğlu (Istanbul: Technical University of Istanbul, 2007), iv1–iv12.
2. See, for example, E. R. Chrastil and W. H. Warren, 'From cognitive maps to cognitive graphs', *PloS One* 9 (2014): e112544.
3. R. A. Doyle, D. Voyer and I. D. Cherney, 'The relation between childhood spatial activities and spatial abilities in adulthood', *Journal of Applied Developmental Psychology* 33 (2012): 112–120; M. Hegarty and D. A. Waller, 'Individual differences in spatial abilities', in *The Cambridge handbook of visuospatial thinking*, edited by A. Miyake and P. Shah (Cambridge: Cambridge University Press, 2005), pp. 121–169; D. Hussain, S. Hanafi, K. Konishi, W. G. Brake and V. D. Bohbot, 'Modulation of spatial and response strategies by phase of the menstrual cycle in women tested in a virtual navigation task', *Psychoneuroendocrinology* 70 (2016): 108–117; T. Ishikawa and D. R. Montello, 'Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places', *Cognitive Psychology* 52 (2006): 93–129; S. C. Levinson, *Space in language and cognition: Explorations in cognitive diversity* (Cambridge, Cambridge University Press, 2003); S. M. Weisberg and N. S. Newcombe, 'Cognitive maps: Some people make them, some people struggle', *Current Directions in Psychological Science* 27 (2018): 220–226; T. Wolbers and M. Hegarty, 'What determines our navigational abilities?' *Trends in Cognitive Sciences* 14 (2010): 138–146.
4. Originally given as the keynote paper to the *Workshop on Space Syntax and Spatial Cognition at the Conference on Spatial Cognition*, 2006, at the University of Bremen.
5. See J. O'Keefe and L. Nadel, *The hippocampus as a cognitive map* (Oxford: Clarendon Press, 1978) – who see in this process the installation of a Kantian *a priori* notion of absolute space.
6. P. Bloom, M. A. Peterson, L. Nadel and M. F. Garrett, 'Space and language', in *Language and space*, edited by M. A. Peterson, L. Nadel, P. Bloom and M. F. Garrett (Cambridge, MA: MIT Press, 1996), p. 569.
7. For example, M. Batty and P. A. Longley, *Fractal cities: A geometry of form and function* (London: Academic Press, 1994); S. Kostof, *The city shaped: Urban patterns and meanings through history* (London: Thames and Hudson, 1992); A. E. J. Morris, *History of urban form, prehistory to the Renaissance* (London: George Godwin Limited, 1972) (later editions published by Prentice-Hall as *History of urban form before the Industrial Revolution*).
8. Editors' note: the original text states that all line maps 'can in small scale cases be created algorithmically by using the DepthMap software'. It then goes on to cite A. Turner, 'DepthMap: A program to perform visibility graph analysis', in *Third International Space Syntax Symposium*, edited by J. Peponis, J. Wineman and S. Bafna (Atlanta, GA: 2001), 31.1–31.9; A. Turner, 'Angular analysis', in *Third International Space Syntax Symposium*, edited by J. Peponis, J. Wineman and S. Bafna (Atlanta, GA: 2001), 30.1–30.11; A. Turner, A. Penn and B. Hillier, 'An algorithmic definition of the axial map', *Environment and Planning B: Planning and Design* 32 (2005): 425–444. The paper then notes that 'for large scale urban systems this is computationally prohibitive, so least line maps are commonly digitised using the rules for creating and checking maps'. Here they cite B. Hillier and A. Penn, 'Rejoinder to Carlo Ratti', *Environment and Planning B: Planning and Design* 31 (2004): 501–511. The editors point out that nowadays analysing large urban systems is no longer an obstacle computationally, and DepthMap software is available open source at https://github.com/SpaceGroupUCL.
9. R. Carvalho and A. Penn, 'Scaling and universality in the micro-structure of urban space', *Physica A: Statistical Mechanics and its Applications* 332 (2004): 539–547; B. Hillier, 'A theory of the city as object: Or, how spatial laws mediate the social construction of urban space', *URBAN DESIGN International* 3–4 (2002): 153–179.
10. B. Hillier, 'The hidden geometry of deformed grids: Or, why space syntax works, when it looks as though it shouldn't', *Environment and Planning B: Planning and Design* 26 (1998): 169–191.
11. B. Hillier, *Space is the machine: A configurational theory of architecture* (Cambridge: Cambridge University Press, 1996).
12. B. Hillier and J. Hanson, *The social logic of space* (Cambridge: Cambridge University Press, 1984).
13. Hillier, 'A theory of the city as object'.
14. B. Hillier and S. Iida, 'Network and psychological effects in urban movement', in *Lecture notes in computer science*, no. 3693, edited by A. G. Cohn and D. M. Mark (Berlin and Heidelberg: Springer, 2005) pp. 475–490; B. Hillier, R. Burdett, J. Peponis and A. Penn, 'Creating life: Or, does architecture determine anything?' *Architecture & Behaviour* 3 (1987): 233–250; B. Hillier, A. Penn, J. Hanson, T. Grajewski and J. Xu, 'Natural movement: Or, configuration and attraction in urban pedestrian movement', *Environment and Planning B: Planning and Design* 20 (1993): 29–66.
15. Hillier and Hanson, *The social logic of space*.
16. Hillier, *Space is the machine*.
17. See note 9 above (eds).
18. For field studies, see Hillier et al., 'Creating life'; for experiments, see R. Conroy, 'Spatial navigation in immersive virtual environments' (PhD diss., University of London, 2001).
19. Hillier, 'A theory of the city as object'.
20. Hillier and Hanson, *The social logic of space*; G. Sabidussi, 'The centrality index of a graph', *Psychometrika* 31 (1966): 581–603.
21. Hillier et al., 'Creating life'; L. C. Freeman, 'A set of measures of centrality based on betweenness', *Sociometry* 40 (1977): 35–41.
22. Hillier and Iida, 'Network and psychological effects in urban movement'.
23. For example, R. C. Conroy, 'The secret is to follow your nose: Route path selection and angularity', *Environment and Behavior* 35 (2003): 107–131; M. Duckham and L. Kulik, '"Simplest" paths: Automated route selection for navigation', in *Lecture notes in computer science*, no. 2825, edited by W. Kuhn, M. F. Worboys and S. Timpf (Berlin: Springer, 2003) pp. 169–185; M. Duckham, L. Kulik and M. Worboys, 'Imprecise navigation', *GeoInformatica* 7 (2003): 79–94; R. G. Golledge, 'Path selection and route preference in human navigation: A progress report', in *International Conference on Spatial Information Theory*, edited by A. U. Frank and W. Kuhn (Berlin: Springer, 1995), pp. 207–222; H. Hochmair and A. U. Frank, 'Influence of estimation errors on wayfinding-decisions in unknown street networks: Analyzing the least-angle strategy', *Spatial Cognition and Computation* 2 (2000): 283–313; Y. O. Kim and A. Penn, 'Linking the spatial syntax of cognitive maps to the spatial syntax of the environment', *Environment and*

Behavior 36 (2004): 483–504; D. R. Montello, *The geometry of environmental knowledge* (Berlin and Heidelberg: Springer, 1992), 136–152; D. R. Montello, *The perception and cognition of environmental distance: Direct sources of information* (Berlin and Heidelberg: Springer, 1997): 297–311; E. K. Sadalla, W. J. Burroughs, and L. J. Staplin, 'Reference points in spatial cognition', *Journal of Experimental Psychology: Human Learning and Memory* 6 (1980): 516–552; S. Timpf, G. S. Volta, D. W. Pollock and M. J. Egenhofer, 'A conceptual model of wayfinding using multiple levels of abstraction', in *Lecture notes in computer science*, no. 639, edited by A. U. Frank, I. Campari and U. Formentini (Berlin: Springer, 1992, 348–367); Turner, 'Depthmap'; S. Winter, 'Modeling costs of turns in route planning', *GeoInformatica* 6 (2002): 345–361.
24. Hillier and Iida, 'Network and psychological effects in urban movement'.
25. Hillier and Iida, 'Network and psychological effects in urban movement'.
26. Hillier and Hanson, *The social logic of space*.
27. O'Keefe and Nadel, *The hippocampus as a cognitive map*. But for a recent review, see L. Nadel and O. Hardt, 'The spatial brain', *Neuropsychology* 18 (2004): 473.
28. G. L. Allen, 'A developmental perspective on the effects of "subdividing" macrospatial experience', *Journal of Experimental Psychology: Human Learning and Memory* 7 (1981): 120.
29. O'Keefe and Nadel, *The hippocampus as a cognitive map*.
30. Hillier and Hanson, *The social logic of space*.

20 The city as a socio-technical system (2012)

Cities as extensions of the human body and mind and potentially the bodies and minds of other species. Introduction to 'The city as a socio-technical system'

Lars Marcus

'The city as a socio-technical system' is one in a long series of articles by Bill Hillier addressing the fundamental humans–environment relation, a recurrent theme throughout his writings.[1] It is important to recognise the fundamental character of the article's subject, which also forms the foundations for space syntax research in general: the humanly constructed spatial configuration of the physical environment. As a field of research and theory this has implications for more or less any human activity, as well as for how we understand humans. Thus, it has relevance for a long series of disciplines.

It is therefore not surprising to find Hillier deliver ideas that resonate with their counterparts in several scientific fields, such as cognition science, complexity science, or, as in this case, information science.[2] We may therefore recognise similarities in how the article is set up and how the argument is made, including its examples and references. However, Hillier always offers a distinct and deep understanding of the context at hand, whereby the original contributions of space syntax to our understanding of the built environment are introduced to shed new light upon the subjects addressed by these sciences.

Even though the subject is most often addressed as the city, it is obvious that the argument is equally relevant to any built environment. The important distinction is rather the ontological divide between a physical and a human city, or in this case, a social and a technical city. This distinction is more original than it may seem, at least when we take into account the rigour of Hillier's ideas. The aim is to identify how the two hang together and the reasons why the humanly constructed city becomes an extension of humans just as other technological devices, such as tools and machines. Hence the idea of the city as a socio-technical system.

In the article, Hillier makes an unusually direct conceptualisation of the city as technology. This in itself may seem surprising, but it goes back to a far broader conception of technology discussed in *Space is the machine*.[3] With reference to anthropology, he makes a distinction there between high-energy specialised *implements* and low-energy generic *facilities* as different forms of technology. Today we tend to only see implements as technology, but Hillier helps us see that human history offers a broader range, including facilities, which include not only buildings but also settlements and cities.

As such, the physical city also plays an essential role in complex urban systems not touched on by Hillier. The physical city is a passive, slow and static system, constructed by bricks and mortar; in scientific terms, it is a simple system. But that does not make it unsophisticated. On the contrary, by being simple and slow and therefore not inclined to surprising behaviour, it comes to control and give order to the complex social and economic systems it accommodates and supports;[4] the physical city endows the city with resilience, a vital quality in all complex systems;[5] what is surprising about cities is not that they are so unpredictable, rather that they are so predictable.

So, the physical city is essential for keeping the unpredictable human city on track. What that track is remains an open question and may vary depending on ideological objectives but also the knowledge of the urban planner. The physical city may embody and thus confer resilience on both undesired ideologies and professional misconceptions. The first question we may leave open for debate; the second is what Hillier aims to remedy by raising the standards and the scientific rigour of our professional understanding of cities.

He does so on a deep level by conceiving how we not only perceive the city with our cognitive apparatus but also create it according to it, so that it becomes an extension of ourselves. But, importantly, it becomes an extension not only of our bodies – so that we can more forcefully physically interact with the environment, but also of our minds – so that we can comprehend it adequately through our perception and cognition. The environment is thereby transformed not only into a socio-technical system but also into what Hillier calls an 'objective subject' – it is a mind-boggling journey he invites us on.

Given new understandings of cities in recent decades, a final thought might be advanced. Is there just one 'objective subject' in the creation of the city? New developments have made us aware that cities also constitute ecosystems. This adds to their complexity. But the subjects in these systems are not only humans but also other species, such as birds and insects, that also interact with and try to make sense of the urban environment. Hence, there may be other 'objective subjects' that we should take into consideration and invite to the co-creation of future cities. If so, Hillier has already explained why.

The city as a socio-technical system

A spatial reformulation in the light of the levels problem and the parallel problem

Bill Hillier

Originally published: Hillier, B. 2012. 'The city as a socio-technical system: A spatial reformulation in the light of the levels problem and the parallel problem'. In: Müller Arisona, S., Aschwanden, G., Halatsch, J. and Wonka, P. (eds), *Digital urban modelling and simulation*. Heidelberg, Dordrecht, London and New York: Springer, 24–48.[6]

Vertical and lateral theories

On the face of it, cities as complex systems are made of (at least) two sub-systems: a physical sub-system, made up of buildings linked by streets, roads and infrastructure; and a human sub-system made up of movement, interaction and activity. As such, cities can be thought of as socio-technical systems. Any reasonable theory of urban complexity would need to link the social and technical sub-systems to each other. Historically, most urban models have sought to make the link through the concept of distance in the physical system as a cost in the social system, defining the physical sub-system as a set of discrete zones. Such models have proved practical tools, but over the years have contributed relatively little to the development of a more general theory of cities. Here we propose a more complex and, we believe, true-to-life model based on the definition of the physical sub-system of the city as a network of spaces – streets and roads – linking buildings, rather than as a system of discrete zones. This allows us to approach urban complexity in a new way.

Two key issues – theoretically perhaps the two key issues – in the study of complexity, in general, are the *levels* problem: how organised complexity, at one level, becomes elementary the next level up; and the *parallel* problem: how systems with different internal dynamics interact with each other. Cohen and Stewart set out a general framework for conceptualising these two problems which is outlined as follows: complex phenomena, at one level, commonly produce lawful (though rarely mathematically describable) emergent simplicities one level up which then have their own emergent dynamic, independent of the complex processes that created them.[7] They call such emergent simplicities, in which 'chaos collapses', 'simplexities'. Simplexities of different kinds then interact and modify each other to create 'complicities', or complexes of simplexities, to construct the complexity of the real world.

Here it is argued that this formulation captures well the problem of complexity in cities as socio-technical systems. To understand cities we need theories that deal with relations across levels, and theories that deal with relations between parallel processes. We might call the former *vertical* theories and the latter *lateral*. A *vertical* theory will be one that works across levels of emergent phenomena, showing how complex distributed processes produce emergent simplicities, which then ignore the complexity of their creation and become independent forces at the emergent level in creating a further level of complexity. A *lateral* theory will then be one that shows how parallel but dynamically independent processes, each with their own emergent simplicities, interact to shape each other.

Singularly clear examples of each can be found in the two most basic processes that create the city: the *vertical* emergence of a network of space from the process of aggregating buildings to create the physical city; and the *lateral* interaction of this emergent pattern with the processes by which different types of economic and social activity locate and organise themselves in space. The *vertical* process, we might say, creates the *spatial* city, the *lateral* process the *functional* city. The outcome of these vertical and lateral processes is the city, which is at once a form of *spatial* patterning and at the same time a form of *social* and *economic* patterning.

These two processes, and the relations between them, are the subjects of this paper. The ultimate focus is on a central question: what is the *mechanism* by which the vertical and lateral connections, on which the operation of the system depends, are made? The answer proposed is the same in both cases: the *human mind*, using intuitive knowledge of *spatial* and *spatio-functional* laws, is the key mechanism. By this, we mean that there exist spatial and spatio-functional laws which, while being wholly objective and describable by simple mathematics, are intuited by people in the same sense that when we throw a ball of paper so that its parabola leads it to land in a waste paper basket we are intuiting the laws of physics. We do so in both cases because these are the laws of the world, and our bodies and minds then learn to operate under the constraints imposed by these laws. An implication of this is that we cannot understand the city without understanding the interaction between human minds and the form of the material world. This is why we must redefine the city as a socio-technical system if we are to develop effective theories about it.

Now this may seem at first a strange pair of ideas: that the emergence of the spatial *form* of the city and the emergence of its pattern of *functioning* both involve laws, and that these laws are somehow imposed on the city through *human cognition*. But both seem to be the unavoidable conclusions of the application of space syntax analysis to study cities over the past two decades. Here we show the research that has led us to these conclusions. First, we show how the use of the space syntax methods to study cities has brought to light a remarkable series of spatial and spatio-functional *regularities* which are pretty well *invariant across cities*, suggesting that there is some kind of *universal city* underlying the diversity of real cities – a remarkable reflection if we bear in mind the heterogeneous social, economic and temporal circumstances in which the world's cities have been created. These regularities are, for us, the theoretical *explicandum* – that which is to be explained – of cities. Second, the use of space syntax as an *experimental tool* has brought to light simple *configurational* laws of space which deal precisely with the effect on emergent patterns of space of different kinds of physical intervention in it, such as placing blocks of buildings in space. We call this the law of *spatial emergence*, and have elsewhere called it the *law of centrality*. This is the law governing the *vertical* process through which the spatial form of the city emerges. Third, the use of space syntax as a tool to study *functional patterns* in cities has brought to light a fundamental relation between the *form* of the city and its *functioning*: that the network of space is, in and of itself, the primary influence on movement flows in different parts of the network, and through this, on emergent patterns of land use and densities, including the shaping of the pattern of centres and subcentres. This is the law of *spatial agency*, and it governs the *lateral* process by which the form of the city shapes its functional patterns. We propose, in effect, that the *regularities* in cities – the *universal city* – can be explained by the *laws*, but that this happens in such a way as to show that *human cognition* is the primary linking mechanism in both processes.

In what follows, we first show how space syntax has brought to light the regularities in the structures of the spatial networks of cities. Then we set out how the laws of the vertical and lateral processes came to light and show how they have shaped the city into a spatio-functional whole. We then show how both processes can be seen in terms of this Cohen-Stewart model, and why this leads to the conclusion that the key mechanism, in both the vertical and lateral processes, is the human mind – not in the sense of real historic individuals, but in the sense of a *generalised individual* acting according to spatial laws which are both objective and intuitively known to humankind in general. We call this generalised, human individual the *objective subject* of the city, meaning, in effect, *all of us*, and how we use our minds in making and using the city. By virtue of being everywhere in space and time in the formation and working of the city, the *objective subject* everywhere imposes its point of view on it, so that cities are cognitive formations in an even more fundamental sense than they are socio-economic formations. The cognitive sets the envelope of possibility within which socio-economic processes create the city. This is why we must revise our definition of the city as *socio-technical* system.

Regularities in urban spatial form

To bring to light the *regularities* in the spatial form of urban networks we must make use of the standard space syntax representation of the network, the *least line map*, or the smallest number of straight lines that cover the system while making all connections.[8]

Examining the least line maps for cities, at all scales and in all parts of the world, we find:

- That at all scales, from local areas to whole cities, cities are made up of a very small number of long lines and a very large number of short lines,[9] so much so that in terms of the line length distributions in their least line maps, cities have been argued to have scale-free properties.[10] This is just as true of more geometric cities such as Chicago and Athens, as it is for more 'organic' (meaning lacking obvious geometry) cities such as Tokyo or London.
- That in 'organic' cities (as defined above), the longer the line, the more likely it is to be end-connected to another by a nearly straight connection (between about 5 and 25 degrees), creating sequences of such lines, which the eye instinctively identifies (see Figures 20.1 and 20.2) when looking at a least line map; the shorter the line, the more likely it is to intersect with others at near right angles, creating local clusters of such lines. In geometrical cities, a similar pattern can be found but with straight, rather than nearly straight, long lines.
- Through these metric and geometric regularities, cities' street networks acquire a *dual* structure, made up of a dominant *foreground network*, marked by linear continuity (and so, in effect, *route continuity*) and a background network, whose more localised character is formed through shorter lines and less linear continuity.

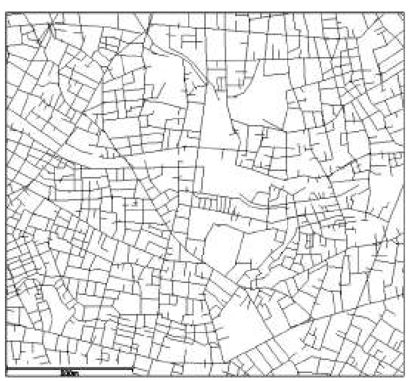


Figure 20.1 Arbitrary section of the least line map of Tokyo.

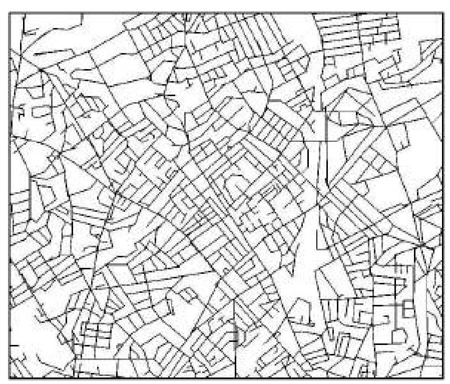


Figure 20.2 Arbitrary section of the least line map of London.

Using the UCL DepthMap software, we can then bring to light configurational regularities in city structures. DepthMap breaks up the least line map into the street segments between junctions and allows three definitions of the distance between each segment and its neighbours: metric, that is the distance in metres between the centre of a segment and the centre of a neighbouring segment; topological, assigning a value of 1 if there is a change of direction between a segment and a neighbouring segment, and 0 if not; and geometric – assigning the degree of the angular change of direction between a segment and a neighbour, so straight connected are 0-valued and a line is a sequence of 0-valued connections. (See Figure 20.3 for an example of an angle-weighted connectivity graph for segment analysis.) It then uses these three concepts of distance to calculate two kinds of measure: syntactic *integration*,[11] which measures how close each segment is to all others under each definition of distance; and syntactic *choice*, or mathematical *betweenness*, which calculates how many distance-minimising paths between every pair of segments each segment lies on, under different definitions of distance. So, using the *metric* definition of distance we find the system of *shortest path* maps for integration and choice, with the *topological* definition we find the system of

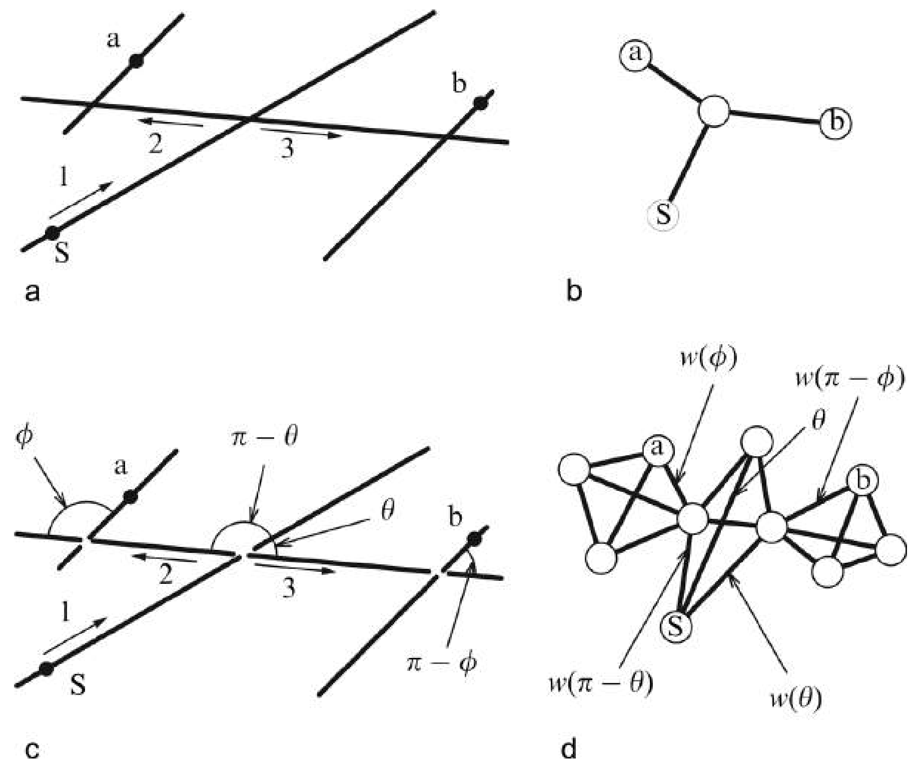


Figure 20.3 Line and segment representation of street networks and their graphs.

fewest-turns maps, and with the *geometrical* definition we find the system of *least angle change* maps. Each of the six measures (two measures with three definitions of distance) can then be applied with the three definitions of distance used as definitions of the *radius* from each segment at which the measures can be applied, giving a total of 18 measures which can, of course, be applied at any radius, so yielding a potentially very large set of possible measures – for example, least angle change choice, at a metric radius of 800 metres – which would be infinite if we count the smallest variation in metric radius.

Applying these measures to cities, we bring to light further regularities. For example:

- By banding mathematical values from dark to light (or in the software from red to blue), meaning from strong to weak, we find, in case after case, least angle integration (normalised closeness) analysis without radius restriction (so the most 'global' form of the analysis) identifies a dominant structure in the form of what we call a *deformed wheel*, meaning a 'hub' of lines in the syntactic centres, strong 'spokes' linking centre to edge and strong 'rim' lines (closely reflecting the patterns brought to light by the earlier syntactic analysis of topological closeness of the least line map). Figures 20.4 and 20.5, for example, show the underlying deformed wheel pattern in both London within the M25 and metropolitan Tokyo (with multiple rims).
- Using the same colouring techniques, the least angle choice (betweenness) measure commonly identifies a network spread through the system, though strongest in the more syntactically central locations (see Figures 20.6 and 20.7).

Figure 20.4 Least angle integration (normalised closeness) for London within the M25.

Figure 20.5 Least angle integration (normalised closeness) for Tokyo.

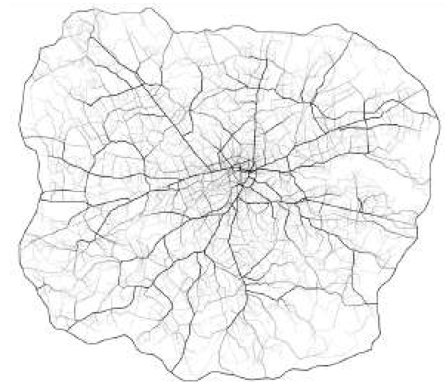


Figure 20.6 Least line angle choice (betweenness) analysis for London.

Figure 20.7 Least line angle choice (betweenness) analysis for Tokyo.

In other words, in spite of the differences in socio-economic and temporal circumstances in which cities grow, they seem to converge on common generic forms, which have metric, geometric and configurational properties in common.

However, the similarities between cities do not stop there. On close examination, for example, all cities seem to exhibit a property we call *pervasive centrality*, meaning that 'central' functions such as retail and catering concentrations diffuse throughout the network at all scales, from the city as a whole to the local network of streets. We typically find that along the length of a high global movement line sequence, the centres occur only in certain locations. For example, if we take the Edgware Road between the North Circular Road and Oxford Street, Figure 20.8, there are three high streets with the rest fairly free of shops. In each case, the centre occurs where local grid intensification (a dense and smaller-scale local grid) coincides with the globally strong alignment (see Figure 20.9 for an illustration of the idea of grid intensification). The pattern is far more complex than envisaged in theories of polycentricity. It is notable also that pervasive centrality seems spatially sustainable because it means that wherever you are, you are close to a small centre and not far from a much larger one.[12] If we then reduce the metric radius of the measures, we find the – much more numerous – smaller-scale centres. For example, at radius 750 metres, all of the 'urban villages' in a section of northwest London are picked out in the darkest shade of grey (Figure 20.10).

These effects are not confined to London or organic grids in general. The same kind of pattern of pervasive centrality was recently found in the historic grid-based city of Suzhou in China. We have also shown it to be the case in Brasília. But it is critical that these effects are found in the least angle map and disappear if we

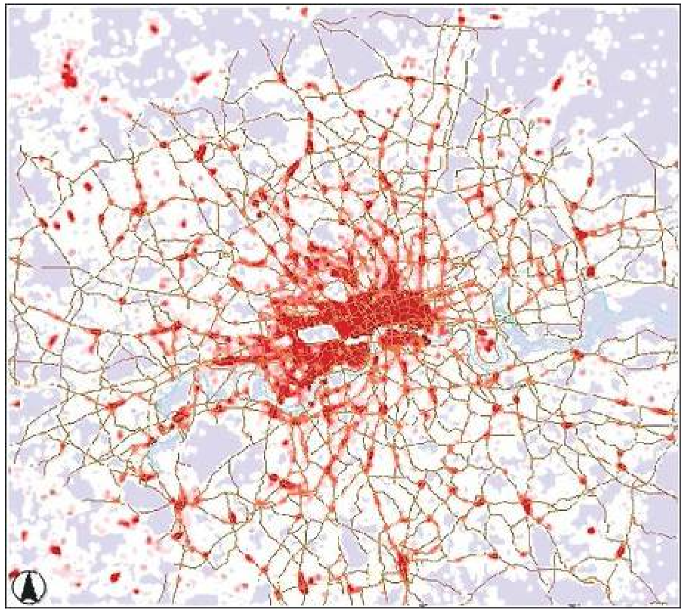


Figure 20.8 Grid intensification coincides with the high street areas in London Image originally published in Batty et al. (2003), Figure 3.[13]

substitute metric for least angle distance in the model. For example, Figure 20.11 shows the pattern of shops in one of the unplanned areas of Jeddah in Saudi Arabia and (inset) the least angle choice measure at a radius of 3.5 kilometres. The match between the two red patterns is remarkable. If we substitute metric for least angle distance (image not shown), we find no relation to the functional patterns. The reasons are simple. In Figure 20.12, we consider three ways of diagonalising a grid. In the top case, the diagonal is regular and so the length of the diagonal route is identical to that of the right side peripheral route. Bottom left, we then create an upward kink on one of the line elements, with the effect of marginally increasing the length of the diagonal route compared to the peripheral route. Bottom right, we create a downward kink on one of line, so marginally shortening the diagonal route compared to the peripheral route, which we show following our usual colouring convention. It follows that with the most marginal changes of this kind, shortest routes will find complex diagonals or simple peripheral routes more or less arbitrarily. This is confirmed in the right figure where we construct a system in which the two diagonals compete, and movement shifts decisively to the downward link and so the shortest path route. Which route is selected by the shortest path algorithm will often then depend on very minor differences in angles, and so be virtually arbitrary.

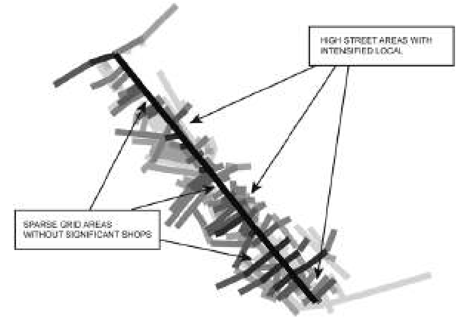


Figure 20.9 Explanation of grid intensification.

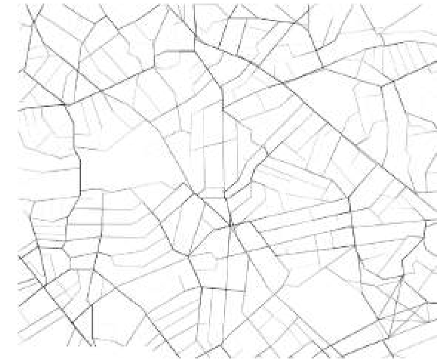


Figure 20.10 Least angle through movement potential at a radius of 750 m in an area of northwest London with the dark lines approximating the urban villages.

Figure 20.11 Least metric distance choice measured at a radius of 3.5 km in Jeddah, Saudi Arabia, with inset showing the pattern of shops in the area marked red.

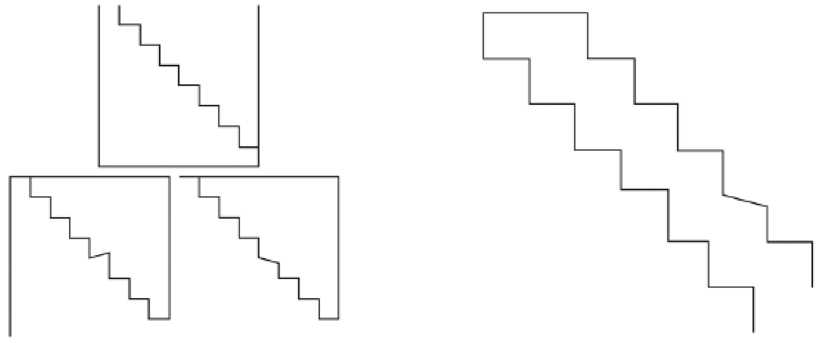


Figure 20.12 Different ways of diagonalising the grid showing why minor geometrical changes can lead to near arbitrary changes in shortest paths.

This arbitrary selection of complex diagonals as shortest paths will feature particularly strongly where a more regular grid system is associated with complex internal structures within grid islands. For example, in Beijing, shortest path choice analysis – Figure 20.12, right – does not find the eight-lane boulevard between the Forbidden City and Tiananmen Square, a boulevard which crosses Beijing east to west and is one of the busiest routes in Beijing. This is then a remarkable failure. In the case of Jeddah, least angle choice analysis, without radius restriction (and so with reference to Jeddah as a whole), picks out the pattern of shops, though more weakly than with the local analysis, Figure 20.13. But substituting metric for least angle distance, we find Figure 20.14, highlighting a nonsense route through the system, with innumerable changes of direction, and with no relation to the functional pattern. At best, we might say that metric analysis helps to identify taxi drivers' routes!

A new definition of the city

The regularities that we find in cities, with least angle analysis, suggest a new definition of the city. Cities of all kinds, however they originate, seem to evolve into *a foreground network of linked centres at all scales, from a couple of shops and a café through to whole sub-cities, set into a background network of largely residential space*. The *foreground* network is made up of a relatively small number of longer lines, connected at their ends by open angles, and forming a super-ordinate structure within which we find the *background* network, made up of much larger numbers of shorter lines, which tend to intersect each other and be connected at their ends by near right angles, and form local grid-like clusters. We suggest this is the proper generic definition of what a city is as a large object.

Figure 20.13 Least angle choice of an area in Jeddah (radius *n*).

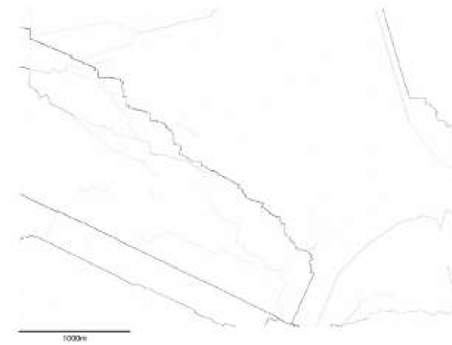


Figure 20.14 Metric choice analysis of an area in Jeddah (radius *n*).

So, what forces give the city this shape? We believe the answer lies in two key new phenomena which research using space syntax has brought to light. The first we call *spatial emergence*: the network of space that links the buildings together into a single system acquires an emergent *structure* from the ways in which objects are placed and shaped within it. This process is law-governed, and without an understanding of these laws, the spatial form of cities cannot really be deciphered. How the city is physically *built* is critical. Cities are not simply reflections of socio-economic processes, but of the *act of building* in the light of these processes. The 'fact of the act' imposes a new framework of lawful constraints on the relation between socio-economic activity and space. It is the *law of spatial emergence* which governs the *vertical* process through which the form of the city's spatial network emerges. The second phenomenon is *spatial agency*: the emergent spatial structure, *in itself*, has lawful effects on the functional patterns of the city by, in the first instance, shaping movement flows, and, through this, emergent land use patterns, since these in their nature either seek or avoid movement flows. Through its influence on movement, the urban grid turns a collection of buildings into a living city. Movement is literally the lifeblood of the city. The *law of spatial agency* governs the *lateral* process through which cities fit functional to spatial patterns.

It is these two linked processes of *spatial emergence* and *spatial agency* that set in train the self-organising processes through which cities acquire their – more or less – *universal* spatial form. These two processes are rendered more or less invisible by the standard method of modelling cities as discrete zones linked by Newtonian attraction. In the syntax approach to network modelling, the differences in attraction found in different parts of the network are *outcomes* of the self-organising process, and so theoretically (as opposed to practically) speaking, should not be taken as a given. But perhaps more than any other factor, it has been the – equally Newtonian! – assumption that space can only be a neutral background to physical processes, rather than an active participant in them, that has rendered these space-based dynamics invisible to urban modelling, and so obscured the path from model to theory. We will now look at *spatial emergence* and *spatial agency* in turn.

A vertical theory: Spatial emergence as a law-governed process

To understand the emergence of the spatial form of the urban network – the vertical problem – we need first to understand its topology then its geometry. The basic form of all cities is one of discrete groups of contiguous buildings, or 'blocks', usually outward facing, defining a network of linear spaces linking the buildings. How can this arise? In fact, very simply.

If we take cell dyads (Figure 20.15, top left), representing buildings linked by entrances to a bit of open space, and aggregate them randomly apart from a rule that each dyad joins its open space cell to one already in the system (forbidding vertex-joins for the buildings, since no one joins buildings corner to corner), a pattern of buildings and spaces emerges with the topology of a city – outward-facing blocks defining a linking network of linear space – but nothing like its geometry, in spite of being constructed on a regular grid.[14] The 'blocks', and so the spaces, are the wrong shape. Where then does the characteristic urban geometry come from?

To understand this, we need first to think a little about the network of space in cities and how we interact with it, and the role that different notions of distance

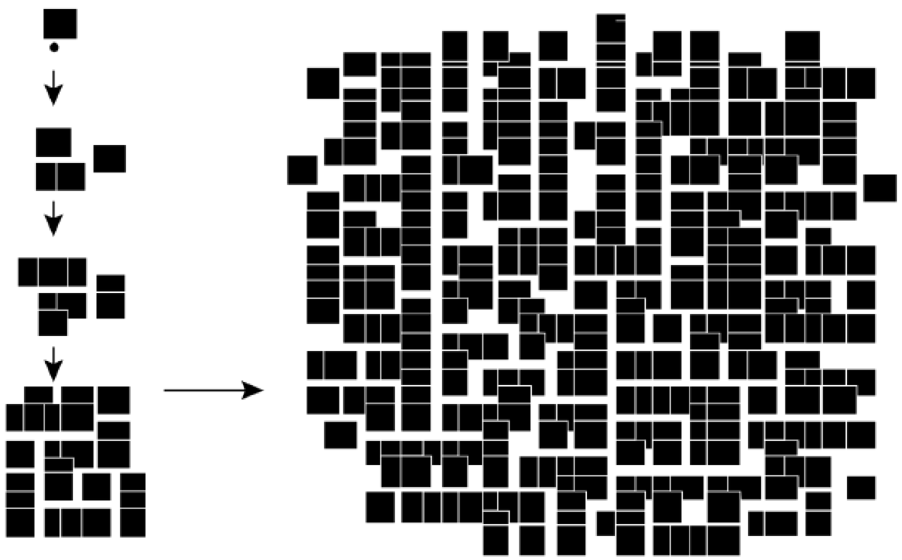


Figure 20.15 Aggregating dyads of open and closed cells by a restricted random process.

might play. Space in cities is about seeing and moving. We interact with space in cities both through our bodies and our minds. Our bodies interact with the space network through moving about in it, and bodily the city exists for us as a system of *metric distances*.

Our minds interact with the city through seeing. By seeing the city, we learn to understand it. This is not just a matter of seeing buildings. We also see space, and the city comes to exist for us also as a visually, more or less complex object, with more or less visual steps required to see all parts from all others, and so as a system of *visual distances*. This warns us that distance in cities might mean more than one thing.

But we also need to reflect on the fact that cities are collective artefacts which bring together and relate very large collections of people. The critical spatial properties of cities are not then just about the relation of one part to another, but of *all parts to all others*. We need a concept of distance which reflects this. We propose that if *specific distance* means the common notion of distance as the distance, visual or metric, from *a* to *b*, that is from an origin to a destination, *universal distance* means the distance from each origin to all possible destinations in the system, and so from all origins to all destinations.[15] Why does this matter? Because universal distance behaves quite differently from the normal metric and geometric concepts of distance that we use habitually. For example, if, as in Figure 20.16, we have to place a cell to block direct movement between two cells, the closer we place it to one of the outer cells, the less the total distance from each cell to all others will be, because more cell-to-cell trips are direct and do not require deviations around the blocking object.

The same applies to inter-visibility from all points to all others, in Figure 20.17. As we move a partition in a line of cells from centre to edge, the total inter-visibility from each cell to all others increases, although, of course, the total area remains constant.

Both metric and visual effects arise from the simple fact that to measure inter-visibility or interaccessibility we need to square the numbers of points on either side of the blockage. So, all we need to know is that twice the square of a number, n, will be a smaller number than $(n-1)^2+(n+1)^2$ and that, in general:

$$2n^2 < (n-x)^2 + (n+x)^2 \quad (1)$$

We can call this the 'squaring law' for space. It applies when, instead of being interested in, say, the distance from *a* to *b*, we are interested in the distance, metric or visual, from each point in the system to all others. In space syntax these 'all-to-all' properties are called *configurational* to distinguish them from simple relational or geometric properties.

So why does this matter? Because how we place and shape physical objects, such as urban blocks, in space, determines the emergent configurational properties of that space. For example, one consequence of the squaring law is that

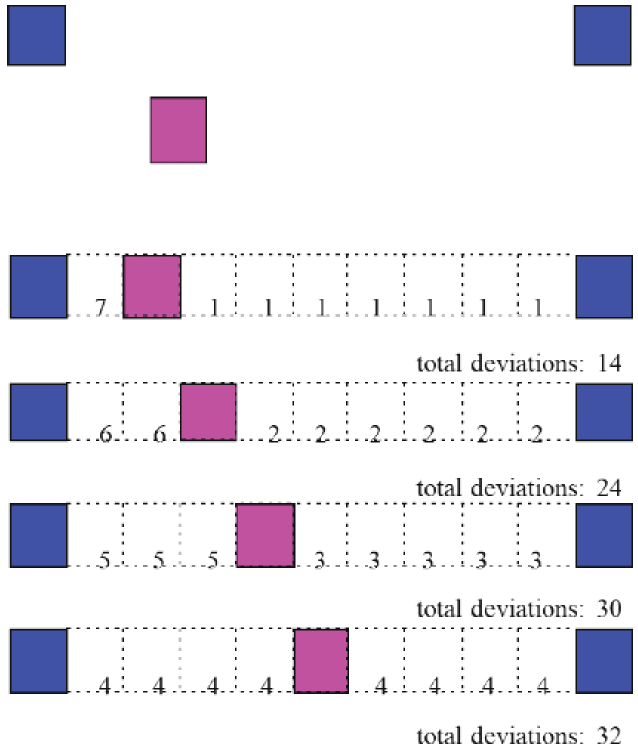


Figure 20.16 Moving an object between two others from edge to centre increases the sum of distances from all cells to all others.

as we move objects from corner to edge and then to central locations in bounded spaces, total inter-visibility in the system decreases, as does visual integration (or universal visual distance), defined as how few visual steps we need to link all points to all others (Figure 20.18a). The same applies to metric integration (or metric universal distance), defined as the sum of the shortest paths between all pairs of points in the ambient space, which increases as we move the obstacle from corner to centre (Figure 20.18b).

The same squaring law governs the effect of shape (Figure 20.19): the more we elongate shapes, keeping the area constant, the more we decrease inter-visibility and increase trip length in the ambient space. The effect of a long and short boundary is to create greater blockage in the system through the squaring law. Even at this stage, this spatial law has a critical implication for cities: in terms of configurational metrics, a short line and a long line are, other things being equal, metrically and visually more efficient in linking the system together than two lines of equal length (Figure 20.20), as would be a large space and a small space, compared to two equal spaces.

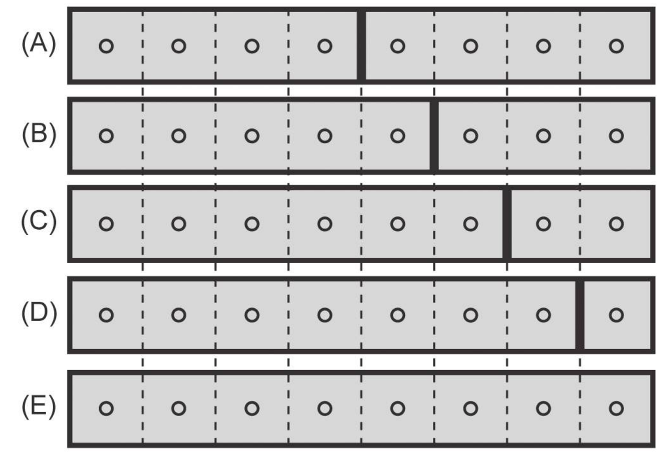


Figure 20.17 Moving a partition from centre to edge increases total inter-visibility. (a) Four points see four cells on both sides, so $2(4)^2 = 32$ or half of the potential for eight cells; (b) five points see five cells on one side, and three see three on the other, so $5^2 + 3^2 = 34$, or 0.53125 of the potential; (c) six see six and two see two, so $6^2 + 2^2 = 40$, or 0.625 of the potential; (d) seven see seven and one sees one, so $7^2 + 1^2 = 50$, or 0.71825 of the potential; (e) eight see eight, so $8^2 + 0^2 = 64$ or all of the potential.

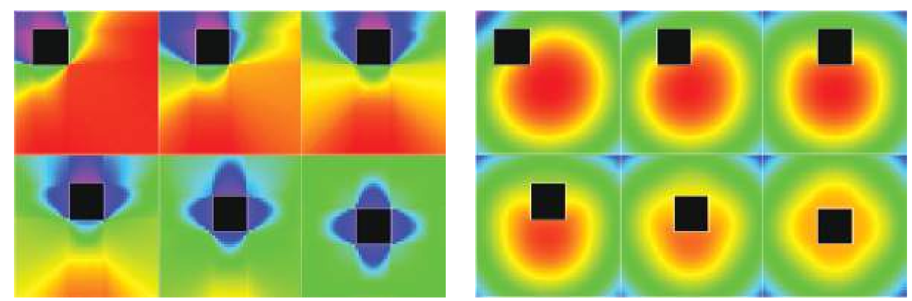


Figure 20.18 (a) Moving an object from corner to centre decreases inter-visibility. Red means less visual distance to all other points, through the spectrum to blue. (b) Moving an object from corner to centre increases the mean length of trips (blue is less metric distance, through the spectrum to red).

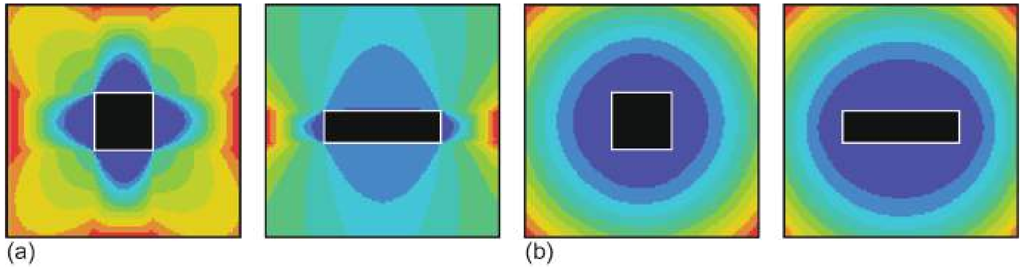


Figure 20.19 Changing the shape of an object from square to rectangular decreases inter-visibility and increases trip length. Colours near the red end of the spectrum from red to blue mean (a) less visual distance and (b) less metric distance.

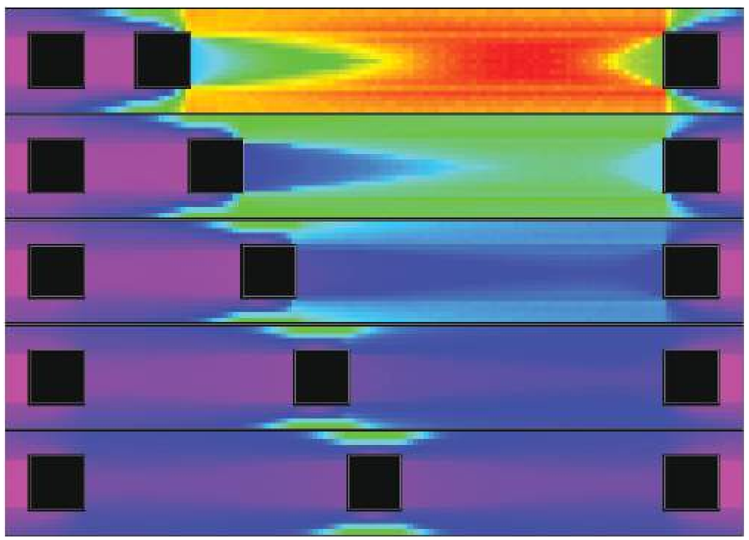


Figure 20.20 Other things being equal, short and long lines integrate more than two lines of equal length. Again, red means less visual distance.

This also has consequences for the mean length of trip (or metric integration) from all points to all others in different types of grid, holding ground coverage of blocks, and therefore total travelable distance in the space, constant. In the four grids in Figure 20.21, blue means shorter mean trip length to all other points. Compared with the regular orthogonal grid (top left), interference in linearity on the right slightly increases mean trip length. But more strikingly, if we reduce the size of central blocks and compensate by increasing the size of peripheral blocks, we reduce mean trip length compared to the regular grid. This of course

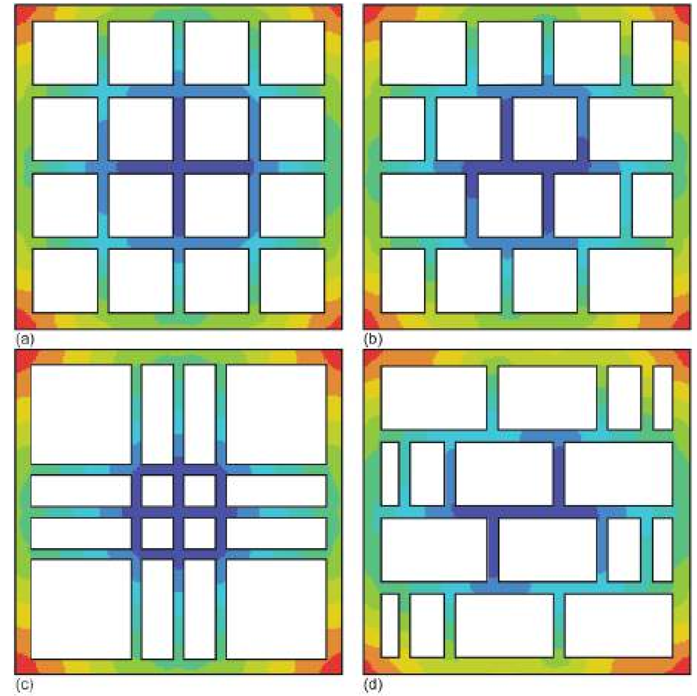


Figure 20.21 Changing the scaling of a grid changes mean trip length. In this case, for graphical clarity, colours near the blue end of the spectrum from red to blue mean less metric distance from each point to all others. The mean distances for each system are: (a) 2.53; (b) 2.59; (c) 2.42; (d) 2.71.

is the ‘grid intensification’ that we often note in looking at centres and subcentres in cities. As so often, we find a mathematical principle underlying an empirical phenomenon.

How we place and shape objects in space then determines the emergent configurational properties of that space. But what kind of block placing and shaping make space urban? On the left of Figure 20.22, we aggregate buildings in an approximately urban way, with linear relations between spaces, so we can see where we are going as well as where we are. On the right we retain the identical blocks but move them slightly to break linear connections between the spaces. If we then analyse metric and visual distances within the two complexes, we find that all-to-all metric distances (not shown) increase in the right-hand case, so trips are on average longer, but the effect is slight compared to the effect on all-to-all visual distances, which changes dramatically (shown in Figure 20.23).

Showing visual integration – dark means less visual distance as before – we see that the left case identifies a kind of main street with side and back streets, so

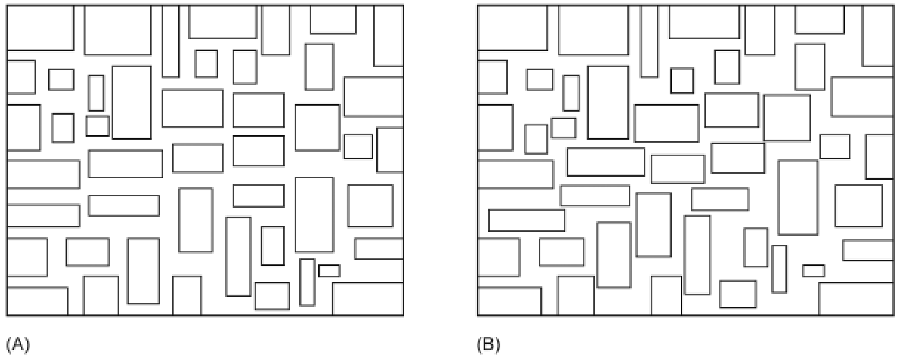


Figure 20.22 Two slightly different arrangements of identical blocks, with strong linear relations between spaces on the left and weak on the right.

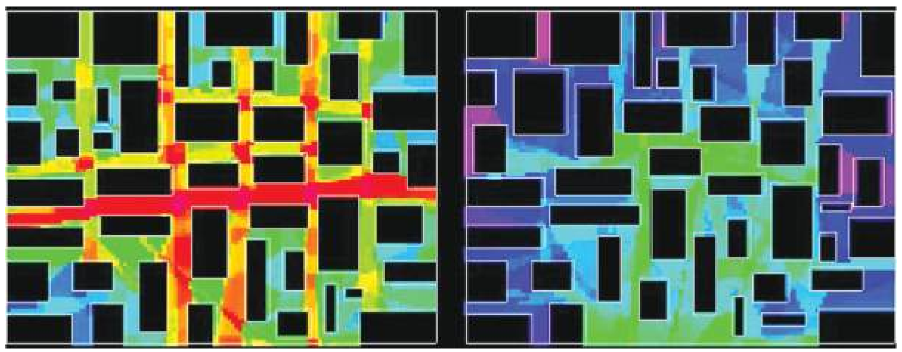


Figure 20.23 Visual integration analysis (red is high, ranging through the temperature spectrum to blue, and so red indicates low visual distances from all points to all others) showing how the non-urban layout (on the right) loses both integration and structure through the slight block changes.

an urban type of structure has emerged. But the right case has lost both structure and degree of inter-visibility. Even though the changes are minor, it feels like a labyrinth. We can see where we are but not where we might be.

The effect on computer agents moving around the system is striking, if obvious. In Figure 20.24, we move 10,000 computer agents with forward vision in the space, again using the software by Alasdair Turner.[16] The agents randomly select a target within their field of vision, move three pixels in that direction, then stop and repeat the process. On the left, the traces of agent movement 'find' the structure of visual integration. On the right, they wander everywhere and tend to get trapped in fatter spaces. This is an effect purely of the configuration, since everything else is identical.

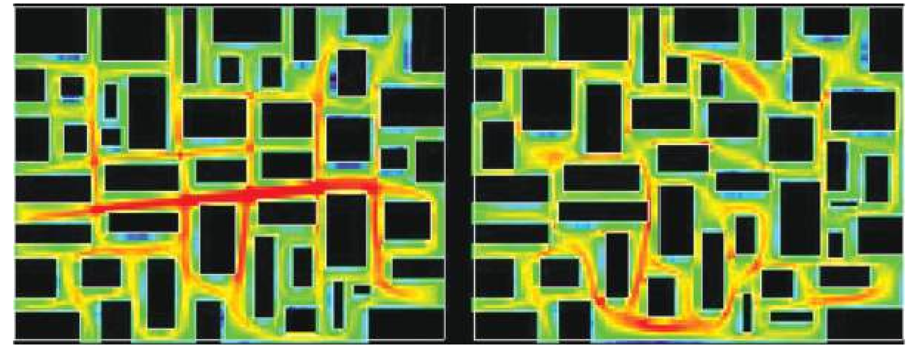


Figure 20.24 Traces of 10,000 forward-looking agents moving nearly randomly in two slightly different configurations. Red means many traces, blue few.

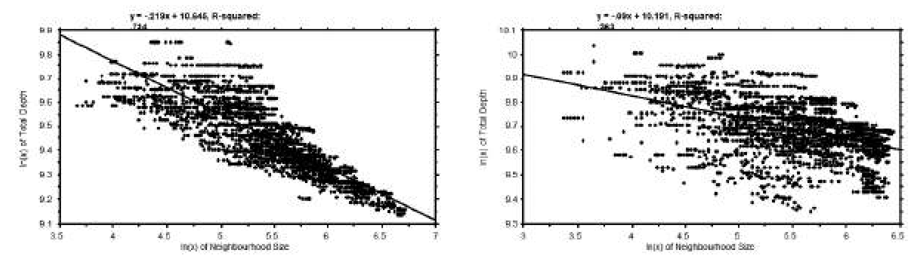


Figure 20.25 Intelligibility scattergrams for the two layouts in Figures 20.13 and 20.14.

But what about human beings? Human beings do not of course move randomly, but purposefully, and successful navigation in an unfamiliar environment would seem to depend on how good a picture of the whole pattern we can get from seeing it from a succession of points within it. One way we might plausibly measure this property is by correlating the size of the visual field we can see from each point with the visual integration value (its visual distance from all others), so in effect measuring the relation between a *local* property that we can see from each point, and a *nonlocal* one that we cannot see (Figure 20.25).

In space syntax this is called the *intelligibility* of the system. The r^2 for the 'intelligible' layout on the left is 0.714 while for the right case it is 0.267. Defined this way, the intelligibility of a spatial network depends almost entirely on its linear structure. Both field studies and experiments suggest that this does work for humans.[17] For example, Conroy Dalton took a linearised 'urban' type network (Figure 20.26, left below) and asked subjects to navigate in a 3D immersive world from left edge to 'Town Square' and back.[18] As the traces show, they manage to find reasonable routes. But she then moved the (identical) blocks slightly to break the linear structure and reduce intelligibility (Figure 20.26, right below) and repeated

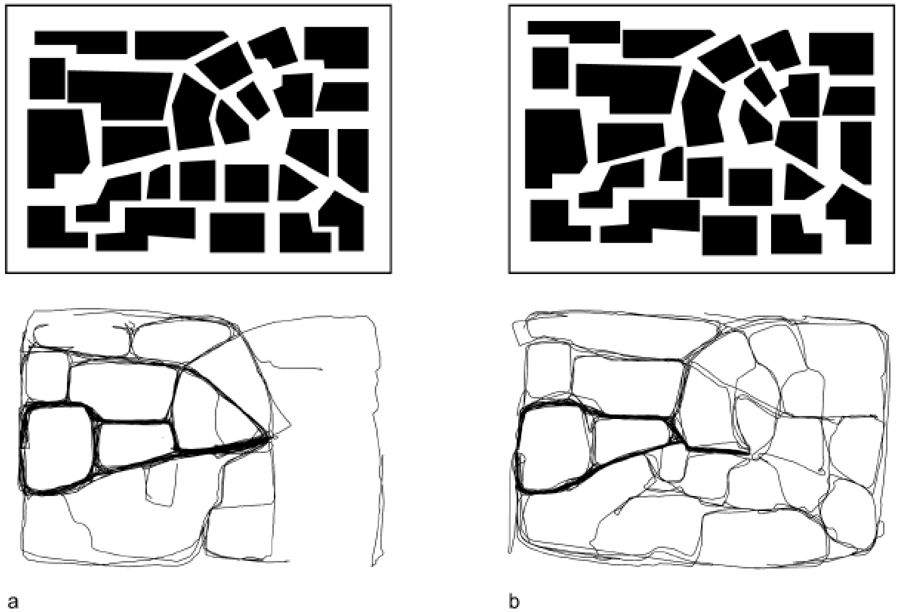


Figure 20.26 Trace of human agents navigating in (a) intelligible and (b) unintelligible layouts. Credit Ruth Conroy Dalton.

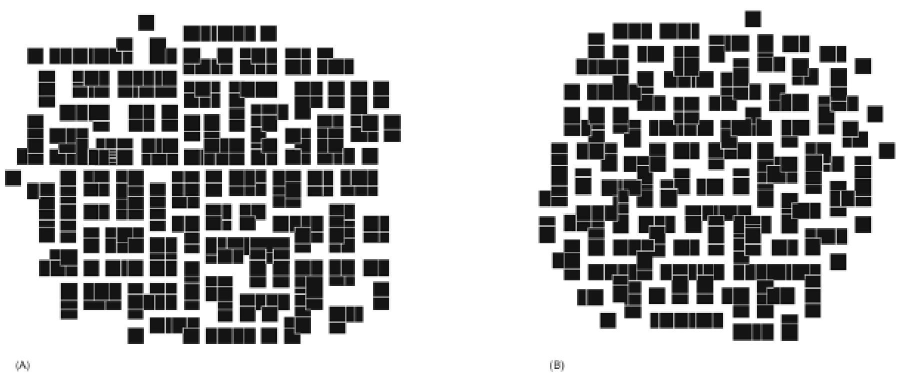


Figure 20.27 (a) A layout generated by a 'conserve longer lines' rule; (b) a layout generated by the inverse rule.

the experiment. The subjects found the modified layout labyrinthine, and many wandered all over the system trying to perform the same wayfinding task.

So if, coming back to our aggregative process, we modify it by requiring those adding cells to the system to avoid blocking a longer local line if they can block a shorter one (Figure 20.27, left) – we might call it a preferential avoidance rule! – we find a layout emerges, which, while still not yet recognisably urban, approximates the mix of long and short lines we find in real systems.[19] With the contrary rule – always block long lines (Figure 20.27, right) – we construct a labyrinth in which lines are of much more even length.

A lateral theory of spatial agency

The vertical process of spatial emergence is then shaped by the squaring law through which the placing and shaping of objects in space creates emergent patterns, and this is why, simply to be intelligible to, and usable by, human beings, spatial networks must include enough long alignments, in proportion to the scale of the settlement itself.[20] The lateral process of spatial agency is then about the consequences of these emergent structures for the functionality of the system. As spatial emergence depends on a spatial law, so spatial agency depends on a spatio-functional law we call the law of natural movement: that other things being equal, the main determinant of movement rates, in different parts of a network, will be a function of the structure of the network itself.[21]

To clarify this, we may first reflect on human movement. Spatially speaking, every human trip is made up of two elements: an origin-destination pair – every trip is from an origin space to a destination space – we can call this the to-movement component; and the spaces passed through on the way from origin to destination – we can call this the through-movement component. It is exactly these two elements of movement, which are captured in the closeness (integration) and betweenness (choice) measures. Integration measures the accessibility of nodes as destinations from origins, so from the principle of distance decay (and other things being equal), we must statistically expect more movement potential for nodes that are closer to all others at some radius. Likewise, since choice measures the sequence of segments we pass through, so we must expect a similar bias in real movement. In effect, integration measures the to-movement, and choice the through-movement, potential of spaces and since we have used these two measures of movement potentials, of both kinds, in urban networks it would be surprising if these potentials did not, to some degree, reflect real movement flows.

But this will depend on how people calculate distances in complex spatial networks, and this is a question much discussed in the cognitive literature.[22] All three measures of distance used in DepthMap – shortest paths, fewest turns, paths and least angle change – have all been canvassed. But we suggest this can be resolved by correlating real flows with the spatial values produced in DepthMap by the three different definitions of distance.[23] Accordingly, we applied the three weightings to the two measures of to- and through-movement potentials to make six different analyses of the same urban system, and correlated the resulting patterns of values for each segment with observed movement flows on that segment (Table 20.1), arguing that if, across cases, there were consistently better correlations with one or other weighting, then the only logical explanation would be that this weighting reflects better how people are biasing spatial movement choices, since everything else about the system is identical. In fact, across four separate studies in areas of central London, we consistently found that geometric, or least angle, weightings yields the strongest movement prediction, with an average of around 0.7 r^2 for vehicular movement and 0.6 r^2 for pedestrian, closely followed by the topological or fewest-turns weighting. Metric shortest paths are markedly inferior in most cases, and in general, to-movement potentials are slightly stronger than through-movement potentials, though this varies from case to case.[24]

Once the law of natural movement is understood, it is clear that the link between the network configuration and movement flows is the key to the *lateral* dynamics and evolution of the system. Because the network shapes movement it also, over time, shapes land use patterns, in that movement-seeking land uses, such as retail, migrate to locations which the network has made movement-rich while others, such as residence, tend to stay at movement-poor locations. This creates multiplier and feedback effects through which the city acquires its universal dual form as a foreground network of linked centres and subcentres at all scales set into a background network of residential space. Through its impact on movement, the network has set in train the self-organising processes by which collections of buildings become living cities.

Table 20.1 Correlations between respective flows and shortest-path, least-angle and fewest-turns analyses, applied to accessibility and choice measures across four areas of central London

a. Vehicular Movement: r^2 values for correlations between vehicular flows and shortest-path, least-angle and fewest-turns analyses, applied to accessibility and choice measures. Best correlations are marked *. Numbers in brackets indicate best radius in segments for accessibility measures.

	Gates	*Measure*	*Least length*	*Least angle*	*Fewest turns*
Barnsbury	116	accessibility	0.131 (60)	0.678 (90)	0.698 (12)
		choice	0.579	0.720*	0.558
Calthorpe	63	accessibility	0.950 (93)	0.837* (90)	0.819 (69)
		choice	0.585	0.773	0.695
South Kensington	87	accessibility	0.175 (93)	0.688 (24)	0.741* (27)
		choice	0.645	0.629	0.649
Brompton	90	accessibility	0.084 (81)	0.692* (33)	0.642 (27)
		choice	0.475	0.651*	0.588

b. Pedestrian Movement: r^2 values for correlations between pedestrian flows and shortest-path, least-angle and fewest-turns analyses, applied to accessibility and choice measures. Best correlations are marked *. Numbers in brackets indicate best radius in segments for accessibility measures.

	Gates	*Measure*	*Least length*	*Least angle*	*Fewest turns*
Barnsbury	117	accessibility	0.119 (57)	0.719* (18)	0.701 (12)
		choice	0.578	0.705	0.566
Calthorpe	63	accessibility	0.061 (102)	0.637 (39)	0.624* (36)
		choice	0.430	0.544	0.353
South Kensington	87	accessibility	0.152 (87)	0.523* (21)	0.502 (15)
		choice	0.314	0.457	0.526*
Brompton	90	accessibility	0.111 (81)	0.623* (63)	0.578 (63)
		choice	0.455	0.513*	0.516

A key element of this will be the formation of centres and subcentres on something like the following lines. Every centre has a centre. Each centre starts with a spatial seed, usually an intersection, but it can be a segment. The seed of a centre will have *destination* and *route* values at both local and global levels. Some – usually small – centres start because they are the focus of a local, intensified grid – a local case – others because they are at an important intersection – a global case. Both global and local properties are relevant to how centres form and evolve. The spatial values of the seed for the centre will establish what we can call a *fading distance* from the seed, which defines the distance from the seed up to which certain land uses, for example shops, will be viable. This is a function of metric distance from the seed proportionate to the strength of the seed. The centre will grow beyond the fading distance established by the initial seed to the degree that further seeds appear within the fading distance, which reinforces the original seed. Again, these can be local or global, and stronger or weaker. A centre becomes larger to the degree that it is reinforced by what are, in effect, new seeds created by the grid, which allow the shopping to be continuous.

Centres then expand in two ways: linearly and convexly. Linear expansion, the most common case, will be along a single alignment or two intersecting alignments, and occurs when the reinforcers are more or less orthogonal or up to 45 degrees to the original alignment or alignments. Convex expansion will be when the shopping streets form a localised grid, and this occurs when reinforcers occur on the parallel as well as the orthogonal alignment. So, centres vary in the strength of their local and global properties and reinforcers, and the balance between them will tend to define the nature of the centre. Most centres will be in some sense strong in both local and global terms, but differences in the balance between local and global will be influential in generating the scale and character of the centre. Centres also grow, or fail, through interaction with neighbouring centres at different scales, and some potential locations for centres fail to be realised due to the existence of a centre close by, but the way in which the urban grid evolves tends to ensure that seeds for potential centres occur only at certain distances from each other.

The dual city of economic and social forces

Building on the *vertical* process, then, the *lateral* process is the means through which economic and social forces put their different imprints on the city. The foreground structure, the network of linked centres, has emerged to maximise grid-induced movement, driven by micro-economic activity. Micro-economic activity takes a universal spatial form and this type of foreground pattern is near-universal in self-organised cities. The residential background network is configured to restrain and structure movement in the image of a particular culture, and so tends to be culturally idiosyncratic, often expressed through a different geometry which makes the city, as a whole look, spatially different. We call the first the *generative* use of space since it aims to generate co-presence and make new things happen, and the second *conservative* since it aims to use space to reinforce existing features of society. In effect, the dual structure has arisen through different effects of the same laws governing the emergence of grid structure and its functional effects. In the foreground space is more random, in the background more rule-governed, so with more conceptual intervention.

We can illustrate this most clearly in a city with more than one culture (now unfortunately separated): Nicosia (Figure 20.28). Top right is the Turkish quarter, bottom left the Greek quarter. Their line geometry is different. In the Turkish quarter, lines are shorter, their angles of incidence have a different range, and there is much less tendency for lines to pass through each other. Syntactically, the Turkish area is much less integrated than the Greek area. We can also show that it is less intelligible and has less synergy between the local and global aspects of space. Yet in spite of these strong cultural differences in the tissue of space, we

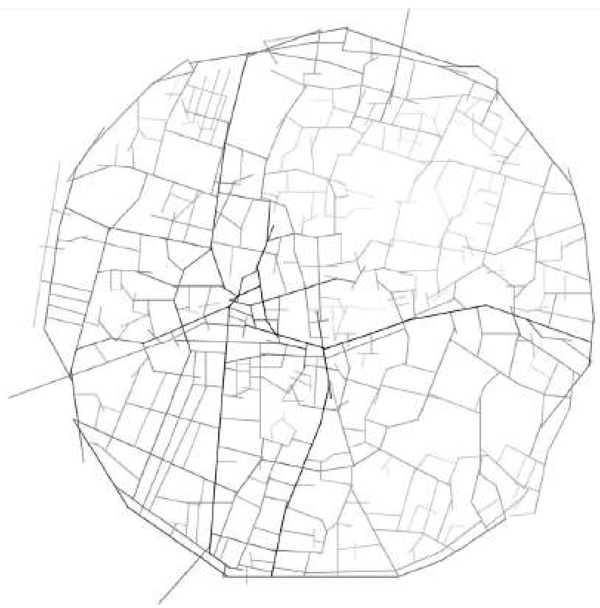


Figure 20.28 The old city of Nicosia with its integration analysis, showing the deformed wheel core in spite of culturally differentiated residential space.

still find Nicosia, as a whole, is held together by a clear deformed wheel structure. This shows how micro-economic activity spatialises itself in a universal way to maximise movement and co-presence, while residence tends to be reflected in the spatial dimension of a particular culture, and the expression is in the first instance geometrical. Since residence is most of what cities are, this 'cultural geometry' tends to dominate our spatial impressions of cities.

The vertical and lateral processes as a simplexity-complicity duo

We see then that the form of the city and its characteristically urban functional patterns emerges from the vertical process of form emergence and the lateral process of function emergence. This is why in self-organised cities, things always seem to be in the right place. It is in the nature of the evolutionary processes through which cities acquire their spatial and functional form. But it is also clear that the vertical and lateral processes form a simplexity-complicity duo in the Cohen-Stewart sense. The low-level, step by step, aggregation of buildings, with the requirement that they be linked to continuous pattern of space, form an emergent pattern of space with its own independent structure – a simplexity – and this independent structure, without regard for the complexities of its creation, then shapes the spatial relation between economic and social processes, which have their own internal dynamics, but which become spatialised in the city through the movement law, so constituting a complicity formed by the interaction of socio-economic and spatial processes.

As we have seen, both the vertical, or simplexity, and the lateral, or complicity, processes are articulated through the intervening medium of spatial and spatio-functional laws. We see this in the fact that only a very small class of spatial forms are created by cities, a vanishingly small proportion of the possible forms that could be constructed with the same raw materials. This is all the more surprising if we bear in mind that almost all large-scale, random aggregates produced by the 'basic generative process' I described are labyrinths, and labyrinths are nowhere found in cities, although we sometimes like to imagine that they are. On the contrary, cities tend towards an improbably high level of integration and intelligibility, and do so through the dual structure, which arises in the first instance from the distribution of line lengths. We can be in no doubt about this, because all cities exhibit dual structures related to the line length distribution, and in all cases the relation to functional processes is of the same kind. Cities are highly improbable forms, but of the same generic kind.

Since the astonishingly tight set of regularities that render cities non-labyrinthine are all expressions of the same basic laws of space, and all the actions that create cities are taken by human beings, it is hard to avoid the inference that the mechanism through which the laws of space reach the spatial form of the city is the human mind, and that this implies that human beings, in general, understand the laws of space. In fact, careful observation of human behaviour easily shows that we do intuitively know the laws of space. Consider this. A group of people is sitting in armchairs in my daughter's flat. My two-year-old grandson comes into the room with two balloons attached to weights by two pieces of string about two and a half feet long, so that the balloons are at about head height for the sitting people. Looking mischievous, he places the balloons in the centre of the space defined by the armchairs. After a minute or two, thinking he has lost interest, one of the adults moves the balloons from the centre of the space to the edge. My grandson, looking even more mischievous, walks over to the balloons and places them back in the centre of the room. Everyone understands intuitively what is going on, including you. But what is actually happening? What my grandson knows is that by placing an object in the centre of a space, he creates more obstruction from all points to all others in that space than if the object is placed anywhere else. In this way, he seeks to draw attention to himself, so that people will interact with him rather than each other. In other words, at the age of two, Freddie knows the laws of space and can use them to achieve a social objective.

Or consider the politics of table shapes. If you take a simple shape, fill it with a fine grid and measure the distance from the centre point of each grid square to the centre point of all other grid squares, the mean distance from central locations to all others is less than those of edge or corner locations. We make the pattern clear by colouring low mean distances in red through to blue for high, using the same range in all cases. As we elongate the object, keeping area constant, we can see that mean distances increase, but the general centre-to-edge pattern is conserved. If we look at the first elongation process, we find that although the overall shape has higher mean distances, the mean distance in the centre of the long side is for a while less than in the centre of the sides in the square shape. This is lost as we elongate more. So, there is a certain point in the elongation of a square to a rectangular shape at which an optimal – in the sense of closer to all other points – edge location is created. It is this simple mathematical fact that is exploited in what we might call the politics of table shapes.

But this evidence of the intervention of human minds, knowledgeable of laws, in the processes of creating cities is circumstantial. We can find much more direct evidence for the intervention of minds in the lateral process through which the grid shapes movement and through this the overall functional patterns of the city. The fact that movement patterns reflect the objective distribution of least angle integration and choice in the system has clear implications. In choosing routes through the urban network, and so in all likelihood estimating distances, people must be using some kind of mental model of the urban grid involving geometric and topological elements, and since urban space can only be experienced as a set of discrete experiences, either of places or routes, these must then be synchronised into a larger-scale pattern for this model to be formed. However, this is exactly the process described by cognitive science in moving from knowledge of routes to map-like knowledge (or 'survey' as they call it). So here we see that this kind of knowledge in human minds, by shaping movement patterns through the network, is shaping the emergent functional patterns in the city itself.

In other words, the key mechanism by which the vertical and lateral processes are linked is actually the human mind itself – in effect, all of us, taking decisions about how to move in the city. This in turn leads to another unavoidable conclusion: that because movement patterns reflect the large-scale, geometrical and topological structure of the network and not the local properties of space, the human mind is actually the means through which cities are created bottom-up by the aggregation of building and spaces, but function top-down through the influence of the larger-scale grid on movement patterns. And the mechanism, by which this remarkable reversal takes place, at the moment when the vertical process first engages the lateral process, is the human mind itself. The vertical-lateral process, which creates the city, is then indecipherable without this knowledge of the intervention of human minds.

The objective subject

The proper theoretical conclusion of these explorations is, I believe, that the human cognitive subject is at the heart of the vertical and lateral processes that create the city, not simply in the sense of a series of real historical individuals located at specific points in time and space, but in terms of the invariance of the cognitive apparatus that those historical individuals bring to the task of creating the city. We are talking, in effect, of a generalised individual located at all points in time and space in the city and everywhere imposing its cognitive apparatus on the ambient city. We might call this generalised individual the *objective subject* of the city-creating process, and therefore of the city.

If, then, it is the case that the city has an objective subject which plays a critical role both in the 'vertical' form-creating process by which the accumulation of built forms creates an emergent spatial pattern, and in the 'lateral' form-function processes by which the emergent spatial pattern shapes movement and sets off the process by which an aggregate of buildings becomes a living city, then what does this imply for our paradigms of the city? The field is broadly split between the social physics paradigm, which seeks to understand the formation of the physical city as the product of spatialised economic processes, and the humanistic or phenomenological paradigm, which seeks to understand the city through our direct experience of it. The social physics view is essentially a mathematical view of the city, while the phenomenological view more or less precludes mathematics. The effect is to create paradigms, which are as irreconcilable methodologically as they are theoretically. The split is made to appear natural by the way we conceptualise our field meta-theoretically as being about the relations between environments simply as material objects and human beings as experiencing 'subjects'.

If the argument in this paper corresponds in any sense to what really happens in cities, then it is clear that environmental 'objects' and human 'subjects' are deeply entangled with each other, with the 'subjective' appearing in the 'objective' world as much as the objective world appears in the human subject. Nor is it the case that the object side of the urban system can be dealt with mathematically and the subject side only qualitatively. The fact that the city is shaped by the human cognitive subject does not lessen its mathematical content since the cognitive processes by which the subject intervenes reflect mathematical laws.[25] We cannot understand the generation of the material form of the city without understanding the formal aspects of the cognitive subject's role in shaping the city, nor understand the experience of the city without knowledge of the formal shape the city acquires under the influence of cognitive subjects.

It follows that we cannot progress while the paradigm split remains. Space syntax was originally created to try to find links between the two previously irreconcilable domains of the city, the city of people and the city of things, hence the 'social' logic of space. The project for space syntax research must now be to engage with the problematic of both the mathematical and humanistic paradigms in the hope and expectation that by finding how each is present in the other we will progress towards synthesis.[26] It could also be instructive for the study of human-mediated complex systems in general, not excluding society itself.[27]

Notes

1. The first publication on this theme being: B. Hillier and A. Leaman, 'The man–environment paradigm and its paradoxes', *Architectural Design* 43 (1973): 507–511.
2. For example, B. Hillier, 'Studying cities to learn about minds: Some possible implications of space syntax for spatial cognition', *Environment and Planning B: Planning and Design* 39 (2012): 12–32; B. Hillier, 'The genetic code for cities: Is it simpler than we think?', in *Complexity theories of cities have come of age: An overview with implications to urban planning and design*, edited by J. Portugali, H. Meyer, E. Stolk and E. Tan (Berlin and Heidelberg: Springer, 2021), pp. 129–152.
3. B. Hillier, *Space is the machine: A configurational theory of architecture* (Cambridge: Cambridge University Press, 1996), p. 180.
4. Compare: W. Weidlich, 'From fast to slow processes in the evolution of urban and regional settlement structures', *Discrete Dynamics in Nature and Society* 3 (1999): 137–147.
5. C. Holling, 'Engineering resilience versus ecological resilience', in *Engineering within ecological constraints*, edited by P. C. Schulze (Washington: National Academies Press, 1996), pp. 31–44.
6. Editors' note: observant readers will note that several passages in this paper repeat sections of the article reproduced in chapter 19 (and published in that same year). We have decided to include it because it develops some of the key ideas outlined in its predecessor, offering an elaboration of the way in which local and global patterns of spatial organisation are nested and interlaced, to create a distributed pattern of centralities at different scales. The distinction of local and global centralities may be traced back to *Space is the machine* (1996) and certainly to 'Centrality as process' (1999 – not included in this volume). This paper is significant both for developing these ideas further, and more clearly than before, and for integrating them into a larger argument regarding the nature of cities as spatial artefacts.
7. J. Cohen and I. Stewart, *The collapse of chaos: Discovering simplicity in a complex world* (Viking: New York, 1994).
8. Editors' note: the original text states that all line maps 'can in small scale cases be created algorithmically by using the UCL DepthMap software'. It then goes on to cite A. Turner, 'DepthMap: A program to perform visibility graph analysis', in *Third International Space Syntax Symposium*, edited by J. Peponis, J. Wineman and S. Bafna (Atlanta, GA: Georgia Institute of Technology, 2001), pp. 31.1–31.9; A. Turner, 'Angular analysis', in *Third International Space Syntax Symposium*, edited by J. Peponis, J. Wineman and S. Bafna (Atlanta, GA: Georgia Institute of Technology, 2001), 30.1–30.11; A. Turner, A. Penn and B. Hillier, 'An algorithmic definition of the axial map', *Environment and Planning B: Planning and Design* 32 (2005): 425–444. The paper then notes that 'for large scale urban systems this is computationally prohibitive, so least line maps are commonly digitised using the rules for creating and checking maps'. Here they cite B. Hillier and A. Penn, 'Rejoinder to Carlo Ratti', *Environment and Planning B: Planning and Design* 31 (2004): 501–511. The editors point out that this is no longer an obstacle computationally, and DepthMap software is available open source at https://github.com/SpaceGroupUCL.
9. B. Hillier, 'A theory of the city as object: Or, how spatial laws mediate the social construction of urban space', *Urban Design International* 3–4 (2002): 153–179.
10. R. Carvalho and A. Penn, 'Scaling and universality in the micro-structure of urban space', *Physica A: Statistical Mechanics and its Applications* 332 (2004): 539–547.
11. This is mathematical closeness with the normalisations set out in B. Hillier and J. Hanson, *The social logic of space* (Cambridge: Cambridge University Press, 1984).
12. B. Hillier, 'Spatial sustainability in cities: Organic patterns and sustainable forms', in *Seventh International Space Syntax Symposium*, edited by D. Koch, L. Marcus and J. Steen (Stockholm: Royal Institute of Technology, 2009), pp. K01:1–K01:20.
13. Batty, M; Besussi, E; Maat, K; Harts, J; (2003) *Representing Multifunctional Cities: Density and Diversity in Space and Time.* (CASA Working Papers 71). Centre for Advanced Spatial Analysis (UCL): London, UK.
14. Hillier and Hanson, *The social logic of space*.
15. Hillier, *Space is the machine*.
16. Turner, 'DepthMap'.
17. See B. Hillier, R. Burdett, J. Peponis and A. Penn, 'Creating life: Or, does architecture determine anything?' *Architecture & Behaviour* 3 (1987): 233–250 for field studies; on experiments, R. Conroy Dalton, 'Spatial navigation in immersive virtual environments' (PhD diss., University of London, 2001) suggests that this does work for humans.
18. R. Conroy, 'Spatial navigation in immersive virtual environments'.
19. Hillier, 'A theory of the city as object'.
20. Hillier, 'A theory of the city as object'.
21. B. Hillier, A. Penn, J. Hanson, T. Grajewski and J. Xu, 'Natural movement: Or, configuration and attraction in urban pedestrian movement', *Environment and Planning B: Planning and Design* 20 (1993): 29–66.
22. For example, S. Winter, 'Modeling costs of turns in route planning', *GeoInformatica* 6 (2002): 345–361; S. Timpf, G. S. Volta, D. W. Pollock and M. J. Egenhofer, 'A conceptual model of wayfinding using multiple levels of abstraction', in *Lecture notes in computer science*, no. 639, edited by A. U. Frank, I. Campari and U. Formentini (Berlin: Springer, 1992), pp. 348–367; H. Hochmair and A. U. Frank, 'Influence of estimation errors on wayfinding-decisions in unknown street networks: Analyzing the least-angle strategy', *Spatial Cognition and Computation* 2 (2000): 283–313; R. Conroy Dalton, 'The secret is to follow your nose: Route path selection and angularity', *Environment and Behavior* 35 (2003): 107–131; M. Duckham and L. Kulik, '"Simplest" paths: Automated route selection for navigation', in *Lecture notes in computer science*, no. 2825, edited by W. Kuhn, M. F. Worboys and S. Timpf (Berlin: Springer, 2003), pp. 165–185; R. G. Golledge, 'Path selection and route preference in human navigation: A progress report', in *International Conference on Spatial Information Theory*, edited by A. U. Frank and W. Kuhn (Berlin: Springer, 1995), pp. 207–222; D. R. Montello, *The geometry of environmental knowledge* (Berlin and Heidelberg: Springer, 1992), pp. 136–152; D. R. Montello, *The perception and cognition of environmental distance: Direct sources of information* (Berlin and Heidelberg: Springer, 1997), pp. 297–311; E. K. Sadalla, W. J. Burroughs and L. Staplin, 'Reference points in spatial cognition', *Journal of Experimental Psychology: Human Learning and Memory* 6 (1980): 516–528; M. Duckham, L. Kulik and M. Worboys, 'Imprecise navigation', *GeoInformatica* 7 (2003): 79–94; Y. O. Kim and A. Penn, 'Linking the spatial syntax of cognitive maps to the spatial syntax of the environment', *Environment and Behavior* 36 (2004): 483–504.
23. B. Hillier and S. Iida, 'Network and psychological effects in urban movement', in *Lecture notes in computer science*, no. 3693, edited by A. G. Cohn and D. M. Mark (Berlin and Heidelberg: Springer, 2005), pp. 475–490.
24. Hillier and Iida, 'Network and psychological effects in urban movement'.
25. Hillier and Iida, 'Network and psychological effects in urban movement'; B. Hillier, 'What do we need to add to a social network to get a society? Answer: something like what we have to add to a spatial network to get a city', *Journal of Space Syntax* 1 (2010): 41–58.
26. B. Hillier, 'Between social physics and phenomenology: Explorations towards an urban synthesis?', in *Fifth International Space Syntax Symposium (2005)*, edited by A. Van Nes (Amsterdam: Techne Press), pp. 3–24.
27. See Hillier, 'What do we need to add'.

References

Adams, R. M. 1966. *The evolution of urban society: Early Mesopotamia and Prehispanic Mexico*. London: Weidenfeld and Nicolson.
Alberti, L. B. 1845 (1452). *De re aedificatoria*, Book IX, *Ornament to Private Buildings*. Translated by James Leoni. Florence: Published by subscription.
Albrecht, G. L., Fitzpatrick, R. and Scrimshaw, S. C. (eds) 2000. *Handbook of social studies in health and medicine*. London: Sage.
Alexander, C. and Ishikawa, S. 1977. *A pattern language*. New York: Oxford University Press.
Allen, G. L. 1981. 'A developmental perspective on the effects of "subdividing" macrospatial experience'. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 120.
Allen, T. 1977. *Managing the flow of technology: Technology transfer and the dissemination of technological information within the R&D organization*. Cambridge, MA: MIT Press.
Anderson, S. (ed.) 1978. *On streets*. Cambridge, MA: MIT Press.
Anon. 1982. 'Report: Hunting Gate amends Limehouse scheme'. *The Architects' Journal*, 175, 42.
Appleton, J. 1984. 'Prospects and refuges re-visited'. *Landscape Journal*, 3, 91–103.
Arbib, M. A. 1969. 'Self-reproducing automata: Some implications for theoretical biology'. In: Waddington, C. H. (ed.), *Towards a Theoretical Biology*. Edinburgh: Edinburgh University Press, 204–226.
Aristotle 1941. *The basic works of Aristotle*. Edited and with an introduction by Richard McKeon. New York: Random House.
Backhouse, A. and Drew, P. 1992. 'The design implications of social interaction in a workplace setting'. *Environment and Planning B*, 19, 573–584.
Bailey, F. G. 1972. *Gifts and poison*. Oxford: Blackwell.
Banham, R. 1960. *Theory and design in the first machine age*. Cambridge, MA: MIT Press.
Barton, K. R., Valtchanov, D. and Ellard, C. 2014. 'Seeing beyond your visual field: The influence of spatial topology and visual field on navigation performance'. *Environment and Behavior*, 46, 507–529.
Batty, M. 2019. 'Foreword: The morphology of cities'. In: D'acci, L. (ed.), *The mathematics of urban morphology*. Cham, Switzerland: Birkhäuser, v–ix.
Batty, M. and Longley, P. A. 1994. *Fractal cities: A geometry of form and function*. London: Academic Press.
Batty, M., Besussi, E., Maat, K., and Harts, J. 2003. *Representing multifunctional cities: Density and diversity in space and time*, CASA Working Papers 71. London: Centre for Advanced Spatial Analysis, UCL.
Benedikt, M. L. 1979. 'To take hold of space: Isovists and isovist fields'. *Environment and Planning B*, 6, 47–65.
Benevolo, L. 1967. *The origins of modern town planning*. Translated [from the Italian] by Judith Landry. London: Routledge and Kegan Paul.
Bernstein, B. 1972. *Class, codes and control*. St Albans: Paladin.
Bevans, J. 1819. 'Design for a penitentiary or gaol for 600 prisoners'. *Reports from Select Committees on Prisons and Gaols 1819–1836 parliamentary papers*. London: House of Commons.
Bloom, P., Peterson, M. A., Nadel, L. and Garrett, M. F. 1996. 'Space and language'. In: Peterson, M. A., Nadel, L., Bloom, P. and Garrett, M. F. (eds), *Language and space*. Cambridge, MA: MIT Press, 553–578.
Boyd, A. 1962. *Chinese architecture and town planning: 1500 B.C.–A.D. 1911*. London: Alec Tiranti.
Bradfield, R. M. 1973. *A natural history of associations: A study in the meaning of community*. London: Duckworth.

Brier, M.-A. and Brunet, P. 1984. *Normandie*. Paris: Berger-Levrault.
Brown, M. G. 1982. 'A theory for the spatial design of organizations'. MSc thesis, University College London.
Buchanan, P. 2007. 'Space syntax and urban design'. In: Jenkins, D. (ed.), *Norman Foster: Works 3.* Munich: Prestel, 178–187.
Buffon, G. 1749. *Histoire naturelle*. Brussels: Lejeune.
Canguilhem, G. 1965. *La connaissance de la vie*. Paris: Vrin.
Carter, H. 1989 (1983). *An introduction to urban historical geography*. London: Edward Arnold.
Carvalho, R. and Penn, A. 2004. 'Scaling and universality in the micro-structure of urban space'. *Physica A: Statistical Mechanics and its Applications*, 332, 539–547.
Cassirer, E. 1944. *An essay on man: An introduction to a philosophy of human culture.* New Haven, CT: Yale University Press.
Chomsky, N. 1957. *Syntactic structures*. The Hague: Mouton.
Chrastil, E. R. and Warren, W. H. 2014. 'From cognitive maps to cognitive graphs'. *PloS One*, 9, e112544.
Coleman, A. 1985. *Utopia on trial: Vision and reality in planned housing*. London: Hilary Shipman.
Collins, P. 1965. 'The linguistic analogy'. In: *Changing Ideals in Modern Architecture 1750–1950.* London and Montreal: McGill University Press, 173–184.
Colquhoun, A. 1967. 'Typology and design method'. *Arena: Journal of the Architectural Association*, 83, 11–14.
Colquhoun, A. 1981. 'Form and figure'. In: *Essays in architectural criticism.* Cambridge, MA: MIT Press, 190–202.
Conroy R. 2001. 'Spatial navigation in immersive virtual environments'. PhD diss., University of London.
Conroy R. 2003. 'The secret is to follow your nose: Route path selection and angularity'. *Environment and Behavior*, 35, 107–131.
Cross, N. 1984. *Developments in design methodology*. Oxford: Wiley-Blackwell.
Cuisenier, J. 1985. 'Type ideal et realites architecturales' (draft of Chapter 1 in mimeo). *La maison rustique: logique sociale et composition architecturale.* Paris: PUF.
Dalton, N. 2001. 'Fractional configurational analysis and a solution to the Manhattan problem'. In: Peponis, J., Wineman, J. and Bafna, S. (eds), *Third International Space Syntax Symposium*. Atlanta, GA: Georgia Institute of Technology, 26.1–26.13.
Descartes, R. 1985 (1644). 'Principles of philosophy'. In: Murdoch, D., Cottingham, J. and Stoothoff, R. (eds), *The philosophical writings of Descartes*. Translated by Dugald Murdoch, John Cottingham and Robert Stoothoff. Cambridge: Cambridge University Press, 177–292.
de Sola Pool, I. and Kochen, M. 1978. 'Contacts and influence'. *Social Networks*, 1, 5–51.
Douglas, M. 1970. *Natural symbols*. London: Barrie and Rockcliffe.
Doyle, R. A., Voyer, D. and Cherney, I. D. 2012. 'The relation between childhood spatial activities and spatial abilities in adulthood'. *Journal of Applied Developmental Psychology*, 33, 112–120.
Dray, W. 1957. *Laws and explanations in history*. Oxford: Oxford University Press.
Duckham, M. and Kulik, L. 2003. '"Simplest" paths: Automated route selection for navigation'. In: Kuhn, W., Worboys, M. F. and Timpf, S. (eds), *Lecture notes in computer science*, no. 2825. Berlin: Springer, 165–185.
Duckham, M., Kulik, L. and Worboys, M. 2003. 'Imprecise navigation'. *GeoInformatica*, 7, 79–94.
Durkheim, É. 1964 (1933). *The division of labour in society*. Translated by George Simpson. New York: The Free Press.
Echenique, M. 1970. 'Models: A discussion'. *Architectural Research and Teaching*, 1, 25–30.
Elman, J. L. 1990. 'Finding structure in time'. *Cognitive Science*, 14, 179–211.
Estienne, C. 1564. *La maison rustique: logique sociale et composition architecturale*. Paris: J. Du Puis.
Ferguson, P., Friedrich, E. and Karimi, K. 2012. 'Origin-destination weighting in agent modelling for pedestrian movement forecasting'. In: Greene, M., Reyes, J. and Castro, A. (eds), *Eighth International Space Syntax Symposium*, Santiago, Chile: Pontificia Universidad Católica de Chile.
Fischer, C. S., Jackson, R. M, Stueve, C. A., Gerson, K., Mccallister Jones, L. and Baldassare, M. 1977. *Networks and places: Social relations in the urban setting*. New York: Free Press.
Fletcher, B. (ed.). 1961. *A history of architecture on the comparative method*, 17th edn, revised by R. A. Cordingley. London: Athlone Press (1st edn 1896).
Fortes, M. 1945. *The dynamics of clanship among the Tallensi: Being the first part of an analysis of the social structure of a trans-Volta tribe by Meyer Fortes*. London: Oxford University Press.
Fortes, M. 1959. *The web of kinship among the Tallensi*. Oxford: Oxford University Press.
Forty, A. 2000. 'Function'. *Words and buildings: A vocabulary of modern architecture.* London: Thames and Hudson.
Foster, Norman. 1997. 'Opening address'. In: Major, M. D., Amorim, L. and Dufaux, F. (eds), *First international space syntax symposium*. London: University College London, i–vi.
Foucault, M. 1970. *The order of things: An archaeology of the human sciences*. New York: Pantheon Books.
Frankl, P. 1973 (1914). *Principles of architectural history: The four phases of architectural style, 1420–1900*. Translated by J. F. O'Gorman. Cambridge, MA, and London: MIT Press.
Fraser, D. 1968. *Village planning in the primitive world*. London: Studio Vista.
Frayling, C. 1994. 'Research in art and design'. *Royal College of Art Research Papers*, 1, 1–5.
Freeman, L. C. 1977. 'A set of measures of centrality based on betweenness'. *Sociometry*, 40, 35–41.
Fromm, E. 1971. 'The significance of the theory of mother right for today'. *The crisis of psychoanalysis*. London: Jonathan Cape.
Gibson, J. 1979. *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
Giddens, A. 1984. *The constitution of society: Outline of the theory of structuration*. Cambridge: Polity Press.
Gil, B. and Coelho, C. 2017. 'Laying the fundamentals: Early methods and intentions from the outsets of space syntax'. *Proceedings of the Eleventh International Space Syntax Symposium*. Instituto Superior Técnico, 13.1–13.13.
Gil, J., Varoudis, T., Karimi, K. and Penn, A. 2015. 'The space syntax toolkit: Integrating depthmapX and exploratory spatial analysis workflows in QGIS'. In: Karimi, K., Vaughan, L., Sailer, K., Palailogou, G. and Bolton, T. (eds), *Proceedings of the Tenth International Space Syntax Symposium*. London: University College London, 108.1–108.18.
Glassie, H. 1975. *Folk housing in middle Virginia: A structural analysis of historic artifacts*. Knoxville: University of Tennessee Press.
GLC 1978. *An introduction to housing layout*. London: Architectural Press.
Gluckman, M. 1960. 'The rise of a Zulu empire'. In: J. G. Jorgensen (ed.), *Biology and culture in modern perspective: Readings from Scientific American.* San Francisco: Freeman, 157–169.
Golledge, R. G. 1995. 'Path selection and route preference in human navigation: A progress report'. In: Frank, A. U. and Kuhn, W. (eds), *International Conference on Spatial Information theory.* Berlin: Springer, 207–222.
Goodman, J. and O'Rourke, J. 1997. *Handbook of discrete and computational geometry*. New York: CRC Press.
Granovetter, M. 1982. 'The strength of weak ties: A network theory revisited'. In: Marsden, P. and Lin, N. (eds), *Social structure and network analysis.* Beverly Hills, CA: Sage, 201–233.
Gutkind, E. A. 1964. *Urban development in Central Europe*. London: Collier-Macmillan.
Hacking, I. 1965. *Logic of statistical inference*. Cambridge: Cambridge University Press.
Hacking, I. 1983. 'The creation of phenomena'. *Representing and intervening: Introductory topics in the philosophy of natural science.* Cambridge: Cambridge University Press, 210–219.
Hall, P. 1998. *Cities in civilization: Culture, innovation, and urban order*. London: Weidenfeld & Nicolson.
Hanson, J. 1989. 'Order and structure in urban space: A morphological history of the City of London'. PhD diss., University of London.
Hanson, J. 1998. *Decoding homes and houses*. Cambridge: Cambridge University Press.
Hanson, J. 2000. 'Urban transformations: A history of design ideas'. *URBAN DESIGN International*, 5, 97–122.
Hanson, J. and Hillier, B. 1979. 'Tradition and change in the English house'. Working paper of the Unit for Architectural Studies, University College London.
Hanson, J. and Hillier, B. 1982. 'Domestic space organisation: Two contemporary space-codes compared'. *Architecture et Comportement/Architecture and Behaviour*, 2, 5–25.
Hanson, J. and Hillier, B. 1987. 'The architecture of community: Some new proposals on the social consequences of architectural and planning decisions'. *Architecture et Comportement/Architecture and Behaviour*, 3, 251–273.
Hanson, J. and Hillier, B. 1993. 'City of London: Shape and role'. *Urban Design Quarterly*, 48, 8–13.
Hegarty, M. and Waller, D. A. 2005. 'Individual differences in spatial abilities'. In: Miyake, A. and Shah, P. (eds), *The Cambridge handbook of visuospatial thinking.* Cambridge: Cambridge University Press, 121–169.
Heidegger, M. 1962 (1927). *Being and time*. Translated by John Macquarrie. Oxford: Blackwell.
Hempel, C. G. 1942. 'The function of general laws in history'. *The Journal of Philosophy*, 39, 35–48.
Hempel, C. G. 1965. *Aspects of scientific explanation*. New York: Free Press.
Hillier, B. 1985. 'The nature of the artificial'. *Geoforum*, 16, 163–178.
Hillier, B. 1985. 'Quite unlike the pleasures of scratching'. *9H*, 7, 66–72.
Hillier, B. 1986. 'Appendix: The distance factor'. *Spatial configuration and use density at the urban level: Towards a predictive model*. Final report to the Science and Engineering Research Council. Unit for Architectural Studies, Bartlett School of Architecture and Planning, University College London.
Hillier, B. 1988. 'Against enclosure'. In: Teymur, N., Markus, T. and Wooley, T. (eds), *Rehumanizing housing.* London: Butterworth, 63–88.
Hillier, B. 1989. 'The architecture of the urban object'. *Ekistics*, 56, 5–21.
Hillier, B. 1991. 'Theory liberates: Intelligent representations and theoretical descriptions in architecture and urban design'. Inaugural lecture presented at University College London, 13 June 1990.

Hillier, B. 1996. 'Cities as movement economies'. *URBAN DESIGN International*, 1, 49–60.
Hillier, B. 1996. *Space is the machine: A configurational theory of architecture*. Cambridge: Cambridge University Press.
Hillier, B. 1998. 'The hidden geometry of deformed grids: Or, why space syntax works, when it looks as though it shouldn't'. *Environment and Planning B: Planning and Design*, 26, 169–191.
Hillier, B. 1999. 'Centrality as a process: Accounting for attraction inequalities in deformed grids'. *URBAN DESIGN International*, 4, 107–127.
Hillier, B. 1999. 'Space as a paradigm for describing emergent structure in strongly relational systems'. In: Holanda, F. de, Amorim, L. and Dufaux, F. (eds), *Second International Space Syntax Symposium*, Brasília. 1–16.
Hillier, B. 2001. 'A theory of the city as object: Or, how spatial laws mediate the social construction of urban space'. In: Peponis, J., Wineman, J. and Bafna, S. (eds), *Third International Space Syntax Symposium*. Atlanta, GA: Georgia Institute of Technology, 2.01–2.28.
Hillier, B. 2002. 'A theory of the city as object: Or, how spatial laws mediate the social construction of urban space'. *URBAN DESIGN International*, 3–4, 153–179.
Hillier, B. 2010. 'What do we need to add to a social network to get a society? Answer: something like what we have to add to a spatial network to get a city'. *Journal of Space Syntax*, 1, 41–58.
Hillier, B. 2012. 'The genetic code for cities: Is it simpler than we think?' In: Portugali, J., Meyer, H., Stolk, E. and Tan, E. (eds), *Complexity theories of cities have come of age: An overview with implications to urban planning and design.* Berlin and Heidelberg: Springer, 129–152.
Hillier, B. 2012. 'Studying cities to learn about minds: Some possible implications of space syntax for spatial cognition'. *Environment and Planning B: Planning and Design*, 39, 12–32.
Hillier, B. 2016. 'The fourth sustainability, creativity: Statistical associations and credible mechanisms'. In: Portugali, J. and Stolk, E. (eds), *Complexity, cognition, urban planning and design.* Berlin and Heidelberg: Springer, 75–92.
Hillier, B., Burdett, R., Peponis, J. and Penn, A. 1987. 'Creating life: Or, does architecture determine anything?' *Architecture et Comportement/Architecture & Behaviour*, 3, 233–250.
Hillier, B., Greene, M. and Desyllas, J. 2000. 'Self-generated neighbourhood consolidation in informal settlements'. *URBAN DESIGN International*, 5, 61–96.
Hillier, B. and Hanson, J. 1982. 'Discovering housing genotypes'. Working paper of the Unit for Architectural Studies, University College London.
Hillier, B. and Hanson, J. 1982. 'Space after modernism'. *9H*, 3, 15–20.
Hillier, B. and Hanson, J. 1984. *The social logic of space*. Cambridge: Cambridge University Press.
Hillier, B., Hanson, J. and Graham, H. 1987. 'Ideas are in things: An application of the space syntax method to discovering house genotypes'. *Environment and Planning B: Planning and Design*, 14, 363–385.
Hillier, B., Hanson, J. and Peponis, J. 1984. 'What do we mean by building function?' In: Powell, J., Cooper, I. and Lera, S. (eds), *Designing for building utilisation.* London: Spon, 61–98.
Hillier, B., Hanson, J., Peponis, J., Hudson, J. and Burdett, R. 1983. 'Space syntax: A different urban perspective'. *The Architects' Journal*, 178(48) 47–63.
Hillier, B. and Iida, S. 2005. 'Network and psychological effects in urban movement'. In: Cohn, A. G. and Mark, D. M. (eds), *Spatial information theory, Lecture notes in computer science*, no. 3693. Berlin and Heidelberg: Springer, 475–490.
Hillier, B. and Leaman, A. 1972. 'A new approach to architectural research'. *RIBA Journal*, 79(12) 517–521.
Hillier, B. and Leaman, A. 1973. 'System, structure and transformation'. *Transactions of the Bartlett Society*, 9, 36–77.
Hillier, B. and Leaman, A. 1973. 'The man–environment paradigm and its paradoxes'. *Architectural Design*, 43, 507–511.
Hillier, B. and Leaman, A. 1974. 'How is design possible?' *Journal of Architectural Research*, 3, 4–11.
Hillier, B. and Leaman, A. 1975. 'The architecture of architecture: Foundations of a mathematical theory of artificial space'. In: Hawkes, D. (ed.), *Models and systems in architecture and building.* Hornby: Construction Press, 5–28.
Hillier, B. and Leaman, A. 1976. 'Architecture as a discipline'. *Journal of Architectural Research*, 5, 28–32.
Hillier, B., Leaman, A., Stansall, P. and Bedford, M. 1976. 'Space syntax'. *Environment and Planning B: Planning and Design*, 3, 147–185.
Hillier, B., Leaman, A., Stansall, P. and Bedford, M. 1978. 'Reply to Professor Leach'. In: Green, D., Haselgrove, C. and Spriggs, M. (eds), *Social organisation and settlement: Contributions from anthropology, archaeology and geography.* Oxford: British Archaeological Reports, 403–405.
Hillier, B., Musgrove, J. and O'Sullivan, P. 1972. 'Knowledge and design'. In: Mitchell, W. (ed.), *AIA-Architect-Researchers' Conference. Environmental design: research and practice. Proceedings of the Third Conference of the Environmental Design Research Association*. Los Angeles: University of California Press, 245–264.

Hillier, B. and O'Sullivan, P. 1989. 'Urban design and climate change'. In: Solar Energy Society (ed.), *Proceedings of Conference C59.* London.
Hillier, B. and Penn, A. 1991. 'Visible colleges: Structure and randomness in the place of discovery'. *Science in Context*, 4, 23–49.
Hillier, B. and Penn, A. 1992. 'Dense civilisations: Or, the shape of cities in the 21st century'. *Applied Energy*, 43, 41–66.
Hillier, B. and Penn, A. 2004. 'Rejoinder to Carlo Ratti'. *Environment and Planning B: Planning and Design*, 31, 501–511.
Hillier, B., Penn, A., Hanson, J., Grajewski, T. and Xu, J. 1993. 'Natural movement: Or, configuration and attraction in urban pedestrian movement'. *Environment and Planning B: Planning and Design*, 20, 29–66.
Hillier, B., Stansall, P. and Hanson, J. 1978. *The analysis of complex buildings*. London: Unit for Architectural Studies, Bartlett School of Architecture and Planning, University College London.
Hillier, B. and Stonor, T. 2010. 'Space syntax: Strategic urban design'. *City Planning Review*, 59, 7–11.
HMSO. 1957. *Conwy Castle and town walls*. London: HMSO.
Hochmair, H. and Frank, A. U. 2000. 'Influence of estimation errors on wayfinding-decisions in unknown street networks – analyzing the least-angle strategy'. *Spatial Cognition and Computation*, 2, 283–313.
Holanda, F. de. 1997. 'Exceptional space'. PhD diss., University of London.
Holling, C. 1996. 'Engineering resilience versus ecological resilience'. In: Schulze, P. C. (ed.), *Engineering within ecological constraints*. Washington, DC: National Academies Press, 31–44.
Horwood, R. 1799. *Plan of the Cities of London and Westminster*. London: Richard Horwood.
Hussain, D., Hanafi, S., Konishi, K., Brake, W. G. and Bohbot, V. D. 2016. 'Modulation of spatial and response strategies by phase of the menstrual cycle in women tested in a virtual navigation task'. *Psychoneuroendocrinology*, 70, 108–117.
Husserl, E. 1981. *Edmund Husserl: Shorter works*. Edited by Peter McCormick and Frederick Elliston. Notre Dame, IN: University of Notre Dame Press.
Irwin, D. 1988. 'The house the architect built'. MSc thesis, University College London.
Ishikawa, T. and Montello, D. R. 2006. 'Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places'. *Cognitive Psychology*, 52, 93–129.
Jacobs, J. 1961. *The death and life of great American cities*. Harmondsworth, Middlesex: Penguin.
Kahneman, D. 2011. *Thinking, fast and slow.* New York: Farrar, Straus and Giroux.
Karimi, K. 1998. 'Continuity and change in old cities: An analytical investigation of the spatial structure in Iranian and English historic cities before and after modernisation'. PhD diss., University of London.
Karimi, K. 2018. 'Space syntax: Consolidation and transformation of an urban research field'. *Journal of Urban Design*, 23, 1–4.
Keene, S. C. 1959. *Introduction to meta-mathematics*. Amsterdam: North-Holland Publishing.
Kelly, G. 1964. *The psychology of personal constructs*. New York: Wiley.
Kim, Y. O. and Penn, A. 2004. 'Linking the spatial syntax of cognitive maps to the spatial syntax of the environment'. *Environment and Behavior*, 36, 483–504.
Kirschenmann, J. C. and Muschalek, C. 1980 (1977). *Residential districts*. Translated [from the German] by TST Translations. London: Harper Collins.
Kostof, S. 1992. *The city shaped: Urban patterns and meanings through history*. London: Thames and Hudson.
Krampen, M. 1979. *Meaning in the urban environment*. London: Pion.
Krier, R. 1979. *Urban space*. London: Academy Editions.
Krüger, M. and Turkienicz, B. 1988. 'Synchrony in urban form'. *Third Seminar on Urban Design in Brazil.* Brasília: University of Brasília.
Krüger, M. 1989. 'On node and axial grid maps: Distance measures and related topics'. *European Conference on the Representation and Management of Urban Change*, Cambridge; working paper version available at https://discovery.ucl.ac.uk/id/eprint/1011/.
Kuhn, T. S. 1962. *The structure of scientific revolutions*. Chicago: University of Chicago Press.
Lakatos, I. 1970. 'History of science and its rational reconstructions'. In: Buck, R. C. and Cohen, R. S. (eds), *PSA 1970: Proceedings of the Biennial Meeting of the Philosophy of Science Association*. Dordrecht: Reidel, 91–136.
Lakatos, I. and Musgrave, A. (eds) 1970. *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press.
Lavedan, P. and Hugueney, J. (eds) 1974. *L'urbanisme au Moyen Age*. Geneva: Bibliotheque de la Societe Francaise d'Archéologie.
Le Corbusier. 1929. 'A contemporary city (La ville contemporaine)'. *The city of tomorrow and its planning.* Translated from the 8th French edition of *Urbanisme* with an introduction by Frederick Etchells. London: John Rodker.

Le Corbusier. 1946 (1927). *Towards a new architecture*. London: Architectural Press.
Leach, E. 1972. 'Anthropological aspects of language: Animal categories and verbal abuse'. In: Maranda, P. (ed.), *Mythology: Selected readings*. Harmondsworth, Middlesex: Penguin Books, 39–68.
Leach, E. 1978. 'Discussion: Does space syntax really "constitute the social"?' In: Green, D., Haselgrove, C. and Spriggs, M. (eds), *Social organisation and settlement: Contributions from anthropology, archaeology and geography*. Oxford: British Archaeological Reports, 385–401.
Lee, R. B. and Devore, I. 1968. *Man the hunter*. Chicago: Aldine.
Lees, L. and Warwick, E. 2022. *Defensible space on the move: Mobilisation in English housing policy and practice*. Newark, NJ: John Wiley & Sons.
Legeby, A. 2013. 'Patterns of co-presence: Spatial configuration and social segregation'. PhD diss., KTH Royal Institute of Technology.
Lévi-Strauss, C. 1966 (1962). *The savage mind*. London: Weidenfeld & Nicholson.
Lévi-Strauss, C. 1972. 'Do dual organizations exist?' In: *Structural Anthropology*. Harmondsworth, Middlesex: The Penguin Press, 132–156.
Levinson, S. C. 2003. *Space in language and cognition: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.
Llewelyn-Davies, R. 1958. 'Deeper knowledge: Better design'. Paper given at Conference on Architectural Education, Magdalen College, Oxford, April 1958.
Loumi, A. 1988. 'Spatial configuration of towns in North Africa'. PhD diss., University of London.
Lyotard, J.-F. 1984. *The postmodern condition: A report on knowledge*. Translated [from the French] by G. Bennington and B. Massumi. Manchester: Manchester University Press.
March, L. 'Models, modes and mores'. 1975. *LUBFS Conference Proceedings: Urban Development Models*. Lancaster: Construction Press, 301–310.
March, L. and Steadman, P. 1971. *The geometry of environment: An introduction to spatial organisation in design*. London: RIBA Publications.
Martin, L. 1972. 'The grid as generator'. In: Martin, L. and March, L. (eds), *Urban space and structures*. Cambridge: Cambridge University Press, 6–27.
Martin, L. and March, L. (eds), 1972. *Urban space and structures*. Cambridge: Cambridge University Press.
Mead, M. 1966. 'The transmission of culture through artefacts'. *Continuities in cultural evolution*. New Haven, CT: Yale University Press, 83–106.
Mellart, J. 1967. *Çatal Hüyük: A Neolithic town in Anatolia*. London: Thames and Hudson.
Michelson, W. H. 1970. *Man and his urban environment: A sociological approach*. Reading, MA: Addison-Wesley Publishing Company.
Michie, D. 1973. 'Machines and the theory of intelligence'. *Nature*, 241, 507–512.
Michie, D. 1974. *On machine intelligence*. Edinburgh: Edinburgh University Press.
Montello, D. R. 1992. 'The geometry of environmental knowledge'. In: Frank, A. U., Campari, I. and Formentini, U. (eds), *Lecture notes in computer science*, no. 639. Berlin and Heidelberg: Springer, 136–152.
Montello, D. R. 1997. 'The perception and cognition of environmental distance: Direct sources of information'. In: Hirtle, S. C. and Frank A. U. (eds), *Lecture notes in computer science*, no. 1329. Berlin and Heidelberg: Springer, 297–311.
Montello, D. R. 2007. 'The contribution of space syntax to a comprehensive theory of environmental psychology'. In: Kubat, A. S., Ertekin, Ö., Güney, Y. I and Eyüboğlu, E. (eds), *Sixth International Space Syntax Symposium*. Istanbul: Technical University of Istanbul, iv1–iv12.
Morris, A. E. J. 1972. *History of urban form, prehistory to the Renaissance*. London: George Godwin Limited.
Moudon, A. V. and Lee, C. 2009. 'Urbanism by numbers: A quantitative approach to urban form'. In: Tatom, J. and Stauber, J. (eds), *Making the metropolitan landscape: Standing firm on the middle ground*. New York: Routledge, 57–77.
Nadel, L. and Hardt, O. 2004. 'The spatial brain'. *Neuropsychology*, 18, 473.
Nagel, E. and Newman, J. R. 1959. *Gödel's proof*. London: Routledge and Kegan Paul.
Newman, O. 1972. *Defensible space: People and design in the violent city*. London: Architectural Press.
Odling-Smee, F. J., Laland, K. N. and Feldman, M. W. 2003. *Niche construction: The neglected process in evolution*. Princeton, NJ: Princeton University Press.
O'Keeffe, J. and Nadel, L. 1978. *The hippocampus as a cognitive map*. Oxford: Clarendon Press.
Panerai, P., Castex, J. and Depaule, J. C. 1975. *Formes urbaines, de l'îlot à la barre*. London: Architectural Press.
Park, G. 1974. *The idea of social structure*. New York: Anchor Books.
Patee, H. H. 1970. 'The problem of biological hierarchy'. In: Waddington, C. H. (ed.), *Towards a theoretical biology*. Edinburgh: Edinburgh University Press, 117–136.
Peet, T. E., Pendlebury, J. D. S., Černý, J., Frankfort, H. and Woolley, L. 1923. *The city of Akhenaten*. London: Egypt Exploration Society.
Penn, A. 2001. 'The role of spatial configuration in spatial cognition: Or, why the axial line?' In: Peponis, J., Wineman, J. and Bafna, S. (eds), *Third International Space Syntax Symposium*. Atlanta, GA: Georgia Institute of Technology, 11.1–11.17.
Peponis, J. 1985. 'The spatial culture of factories'. *Human Relations*, 38, 357–390.
Peponis, J., Hadjinikolaou, E., Livieratos, C. and Fatouros, D. A. 1989. 'The spatial core of urban culture'. *Ekistics*, 56, 43–55.
Peponis, J., Ross, C. and Rashid, M. 1997. 'The structure of urban space, movement and co-presence: The case of Atlanta'. *Geoforum*, 28, 341–358.
Peponis, J., Wineman, J., Bafna, S., Rashid, M. and Kim, S. H. 1998. 'On the generation of linear representations of spatial configuration'. *Environment and Planning B: Planning and Design*, 25, 559–576.
Peponis, J., Wineman, J., Rashid, M., Kim, S. H. and Bafna, S. 1997. 'On the description of shape and spatial configuration inside buildings: Convex partitions and their local properties'. *Environment and Planning B: Planning and Design*, 24, 761–781.
Pérez-Gómez, A. 1983. *Architecture and the crisis of modern science*. Cambridge, MA: MIT Press.
Piaget, J. 1968. 'Developmental psychology'. In: Sills, D. L. (ed.), *International encyclopedia of the social sciences*. New York: Macmillan, 136–147.
Piaget, J. 1969. *The mechanisms of perception*. Translated [from the French] by G. N. Seagrim. London: Routledge and Kegan Paul.
Piaget, J. 1971. *Structuralism*. Translated [from the French] and edited by C. Maschler. London: Routledge and Kegan Paul.
Piaget, J. 1971. *Biology and knowledge: An essay on the relations between organic regulations and cognitive processes*. Translated [from the French] by B. Walsh. Edinburgh: Edinburgh University Press.
Piaget, J. and Inhelder, B. 1956. *The child's conception of space*. London: Routledge and Kegan Paul.
Plato. 1941. *The Republic of Plato*. Translated and edited by F. MacDonald Cornford. London: Oxford University Press.
Plato, Cairns, H. and Hamilton, E. 1963. *The collected dialogues of Plato including the letters*. Edited by E. Hamilton and H. Cairns. Princeton, NJ: Princeton University Press.
Plato. 1970. *The dialogues of Plato: Philebus*. Translated and introduced by B. Jowett. London: Sphere Books.
Polanyi, K. 1968. 'Aristotle discovers the economy'. In: Dalton, G. (ed.), *Primitive, Archaic and Modern Economies: Essays of Karl Polanyi*. London: Anchor Books, 78–115.
Pollio, M. V. 1874. *The architecture of Marcus Vitruvius Pollio, in ten books*. London: Lockwood.
Popper, K. 1959. *The logic of scientific discovery*. London: Hutchinson.
Popper, K. 1963. *Conjectures and refutations: The growth of scientific knowledge*. London: Routledge and Kegan Paul.
Popper, K. 1968. 'Epistemology without a knowing subject'. In: Van Rootselaar, B. and Staal, J. F. (eds), *Studies in Logic and the Foundations of Mathematics*. Amsterdam: North-Holland Publishing, 333–373.
Prussin, L. 1969. *Architecture in Northern Ghana*. Berkeley: University of California Press.
Pushkarev, B. and Zupan, J. M. 1975. *Urban space for pedestrians*. Cambridge, MA: MIT Press.
Ravetz, A. 1974. *Model estate: Planned housing at Quarry Hill*. Leeds: Croom Helm, in association with the Joseph Rowntree Memorial Trust.
Raw, G. J. 1992. *Sick building syndrome: A review of the evidence on causes and solutions*. London: HMSO.
Reed, E. K. 1956. 'Types of village-plan layouts in the Southwest'. In: Willey, G. R. (ed.), *Settlement Patterns in the New World*. New York: Wenner-Gren Foundation, 15–16.
Richards, I. A. 1963. 'How does a poem know when it is finished?' In: Lerner, D. (ed.), *Parts and wholes*. New York: Free Press of Glencoe, 163–174.
Roberts, F. H. H. 1931. 'Ruins at Kiatuthlanna Eastern Arizona'. *Bureau of American Ethnology Bulletin*, 100. Washington, DC: Smithsonian Institution.
Rossi, A. 1985 (1966). *The architecture of the city*. Cambridge, MA: MIT Press.
Rowe, C. 1976. *The mathematics of the ideal villa and other essays*. Cambridge, MA: MIT Press.
Royal Institute of British Architects. 1991. *A research policy for the profession*. London: Royal Institute of British Architects.
Royal Institute of British Architects Research Committee. 1970. *Strategies for architectural research: Architectural research and teaching*. London: Royal Institute of British Architects.
Rudovsky, B. 1964. *Architecture without architects*. London: Academy Editions.
Russell, B. 1911. *The problems of philosophy*. Oxford: Oxford University Press.
Rykwert, J. 1976. *The idea of a town: The anthropology of urban form in Rome, Italy and the ancient world*. Princeton, NJ: Princeton University Press.
Sabidussi, G. 1966. 'The centrality index of a graph'. *Psychometrika*, 31, 581–603.
Sadalla, E. K., Burroughs, W. J. and Staplin, L. J. 1980. 'Reference points in spatial cognition'. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 516–528.
Sahlins, M. 1974. *Stone Age economics*. London: Tavistock Publications.

Salah-Salah, F. 1987. 'Cities in the Sahara: Spatial structure and generative processes'. PhD diss., University of London.
Sanoff, H. 1977. *Methods of architectural programming*. Stroudsbourg, PA: Dowden, Hutchinson & Ross.
Santos, C. N. F. 1981. *Movimentos urbanos no Rio de Janeiro*. Rio de Janeiro: Zahar.
Sassen, S. 1991. *The global city*. Princeton, NJ: Princeton University Press.
Saussure, F. de. 1966 (1959). *Course in general linguistics*. New York: McGraw Hill.
Schön, D. A. 1987. *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. San Francisco: Jossey-Bass.
Schön, D. A. 1992. *The reflective practitioner: How professionals think in action*. New York: Basic Books.
Scruton, R. 1977. *The aesthetics of architecture*. London: Methuen.
Scruton, R. 1981. *A short history of modern philosophy: From Descartes to Wittgenstein*. London: ARK Paperbacks.
SD. 1934. *Sederon numbers 5–6, Saint Trinit, Section E, 2e Feuille*. France: Service de la Documentation Nationale du Cadastre.
Serlio, S. 1566. *D'architettura*. Venice: Senese and Krugher, 2:25.
Service, E. R. 1971. *Primitive social organization: An evolutionary perspective*. New York: Random House.
Shannon, C. and Weaver, W. 1948. *The mathematical theory of communication*. Chicago and London: Illinois University Press.
Simon, H. 1969. *The sciences of the artificial*. Cambridge, MA, and London: MIT Press.
Skaburskis, J. V. 1974. Territoriality and its relevance to neighbourhood design: a review. *Journal of Architectural Research*, 3(1), 39–44.
Spencer, H. 1877. *The principles of sociology*. London: Williams and Norgate.
Steadman, P. 1983. *Architectural morphology*. London: Pion.
Stubbs, S. A. 1950. *Bird's-eye view of the Pueblos*. Norman: University of Oklahoma Press.
Tarn, J. N. 1973. *Five per cent philanthropy: An account of housing in urban areas between 1840 and 1914*. Cambridge: Cambridge University Press.
Thayer, H. S. (ed.) 1953. *Newton's philosophy of nature: Selections from his writings*. New York: Hafner Publishing Company.
Thom, R. 1968. 'Comment by René Thom'. In: Waddington, C. H. (ed.), *Towards a theoretical biology*. Edinburgh: Edinburgh University Press, 32–41.
Thom, R. 1969. 'Topological models in biology'. *Topology*, 8, 313–335.
Thom, R. 1970. 'Topologie et linguistique'. *Essays on topology and related topics*. Berlin and Heidelberg: Springer.
Thom, R. 1970. 'Topological models in biology'. In: Waddington, C. H. (ed.), *Towards a theoretical biology*. Edinburgh: Edinburgh University Press, 89–116.
Thom, R. 1972. 'Structuralism and biology'. In: Waddington, C. H. (ed.), *Towards a theoretical biology*. Edinburgh: Edinburgh University Press, 68–82.
Thom, R. 1974. *Modèles mathématiques de la morphogenèse: recueil de textes sur la théorie des catastrophes et ses applications*. Paris: 10/18.
Thom, R. 1975. *Structural stability and morphogenesis: An outline of a general theory of models*. Reading, MA: Benjamin.
Timpf, S., Volta, G. S., Pollock, D. W. and Egenhofer, M. J. 1992. 'A conceptual model of wayfinding using multiple levels of abstraction'. In: Frank, A. U., Campari, I. and Formentini, U. (eds), *notes in computer science*, no. 639. Berlin: Springer, 348–367.
Trnka, B. 1970. 'Prague structural linguistics'. In: Lane, M. (ed.), *Structuralism: A reader*. London: Jonathan Cape, 73–84.
Tsoskounoglou, E. 1995. 'Spatial vulnerability to crime in the design of housing estates'. PhD diss., University of London.
Turner, A. 2001. 'Angular analysis'. In: Peponis, J., Wineman, J. and Bafna, S. (eds), *Third International Space Syntax Symposium*. Atlanta, GA: Georgia Institute of Technology, 30.1–30.11.
Turner, A. 2001. 'DepthMap: A program to perform visibility graph analysis'. In: Peponis, J., Wineman, J. and Bafna, S. (eds), *Third International Space Syntax Symposium*. Atlanta, GA: Georgia Institute of Technology, 31.1–31.9.
Turner, A., Doxa, M., O'Sullivan, D. and Penn, A. 2001. 'From isovists to visibility graphs: A methodology for the analysis of architectural space'. *Environment and Planning B: Planning and Design*, 28, 103–121.
Turner, A. and Penn, A. 2002. 'Encoding natural movement as an agent-based system: An investigation into human pedestrian behaviour in the built environment'. *Environment and Planning B: Planning and Design*, 29, 473–490.
Turner, A., Penn, A. and Hillier, B. 2005. 'An algorithmic definition of the axial map'. *Environment and Planning B: Planning and Design*, 32, 425–444.
Turner, V. 1969. *The ritual process: Structure and anti-structure*. New York: Cornell University Press.
Turner, V. W. 1957. *Schism and continuity in an African society*. Manchester: Manchester University Press.
Vaughan, L., Jones, C. E., Griffiths, S. and Haklay, M. 2010. 'The spatial signature of suburban town centres'. *Journal of Space Syntax*, 1, 77–91.
Venturi, R. 1966. *Complexity and contradiction in architecture*. New York: Museum of Modern Art.
Vitruvius 1960. *The ten books on architecture*. Translated by M. Hicky Morgan. New York: Dover Publications.
Vogt, E. Z. 1968. 'Some aspects of Zinacantan settlement patterns and ceremonial organization'. In: Chang, K. (ed.), *Settlement archaeology*. Palo Alto, CA: National Press Books, 154–173.
Webber, M. M. (ed.) 1964. *Explorations into urban structure*. Philadelphia: University of Pennsylvania Press.
Weber, M. 1947. *The theory of social and economic organisation*. Glencoe, IL: Free Press.
Weber, M. 1958 (1921). *The city*. Translated and edited by D. Martindale and G. Neuwirth. New York: Free Press.
Weidlich, W. 1999. 'From fast to slow processes in the evolution of urban and regional settlement structures'. *Discrete Dynamics in Nature and Society*, 3, 137–147.
Weisberg, S. M. and Newcombe, N. S. 2018. 'Cognitive maps: Some people make them, some people struggle'. *Current Directions in Psychological Science*, 27, 220–226.
Weyl, H. 1963. *Philosophy of mathematics and natural science*. New York: Atheneum.
Wheatley, P. 1972. 'The concept of urbanism'. In: Ucko, P. J., Tringham, R. and Dimbleby, G. W. (eds), *Man, settlement and urbanism: Proceedings of a meeting of the Research Seminar in Archaeology and Related Subjects held at the Institute of Archaeology, London University*. London: Duckworth, 601–637.
Whitehead, A. N. 1961 (1941). 'Mathematics and the good'. In: Johnson, A. H. (ed.), *The interpretation of science: Selected essays*. Indianapolis: Bobbs-Merrill, 187–203.
Willmott, R. 1997. 'Structure, culture and agency: Rejecting the current orthodoxy of organisation theory'. *Journal for the Theory of Social Behaviour*, 27, 93–123.
Wimsatt, W. K. and Beardsley, M. C. 1946. 'The intentional fallacy'. *The Sewanee Review*, 54, 468–488.
Winograd, T. 1972. *Understanding natural languages*. Cambridge: Cambridge University Press.
Winter, S. 2002. 'Modeling costs of turns in route planning'. *GeoInformatica*, 6, 345–361.
Wirth, L. 1938. 'Urbanism as a way of life'. *The American Journal of Sociology*, 44, 1–24.
Wittgenstein, L. 1922. *Tractatus logico-philosophicus*. London: Kegan Paul.
Wittgenstein, L. 1953. *Philosophical investigations*. Oxford: Blackwell.
Wolbers, T. and Hegarty, M. 2010. 'What determines our navigational abilities?' *Trends in Cognitive Sciences*, 14, 138–146.
Wolf, E. R. 1966. *Peasants*. Englewood Cliffs, NJ: Prentice-Hall.
Wood-Jones, R. B. 1963. *Traditional domestic architecture of the Banbury region*. Manchester: Manchester University Press.
Young, M. and Willmott, P. 1962 (1957). *Family and kinship in East London*. Harmondsworth, Middlesex: Penguin.
Zeliotis, C. 1985. 'The spatial structure of Venice'. MSc thesis, University College London.
Zipf, G. K. 1949. *Human behavior and the principle of least effort: An introduction to human ecology*. Cambridge, MA: Addison-Wellesley Press.

Contributors' biographies

Ordered alphabetically by surname.

Luiz Amorim is Professor of Architecture at the Federal University of Pernambuco. He served as a part-time member of the Federal University of Alagoas (1983–1986), invited professor at Taubman College, University of Michigan (2004), and researcher at the Instituto Superior Técnico, University of Lisbon (2017–2018). He is currently Professor at the Federal University of Paraiba. Bill Hillier supervised his doctoral studies at the Bartlett between 1994 and 1999.

Sonit Bafna is an Associate Professor at the School of Architecture, Georgia Institute of Technology, and the Director of its PhD Program. He trained as an architect at CEPT, Ahmedabad, and, after an MSc in Architectural Studies from MIT, received a PhD from Georgia Tech for work in space syntax and historical interpretation. He has served as a faculty member at Georgia Tech since 2002.

Meta Berghauser Pont is Professor in Urban Morphology and Urban Design at Chalmers University of Technology in Gothenburg, Sweden. She developed the Spacematrix method together with Per Haupt to quantitatively describe building types. In 2010, she joined the Spatial Analysis and Design group, led by Lars Marcus, where they integrated Spacematrix with approaches central to space syntax. Since 2023, she has been leading the Spatial Morphology Group (SMoG) at Chalmers.

Ricky Burdett is Professor of Urban Studies and Director of LSE Cities at the London School of Economics and Political Science. He has acted as an adviser to national and metropolitan governments – including the Mayor of London – as well as private organisations on major urban regeneration projects. Burdett received an MSc in Advanced Architectural Studies from the Bartlett and worked as a researcher in the unit directed by Bill Hillier in the early to mid-1980s.

Ruth Conroy Dalton has been Professor of Architecture at the University of Northumbria at Newcastle since 2010. She studied for her BSc, MSc and PhD at University College London, the latter supervised by Bill Hillier (2001). Her first academic appointment was at the Georgia Institute of Technology, 2001–2004, after which she returned to UCL where she served briefly, 2007–2009, as Joint-Programme Director, MSc Advanced Architectural Studies (now MSc Space Syntax).

Sam Griffiths is Associate Professor in Spatial Cultures at the Space Syntax Laboratory, Bartlett School of Architecture, University College London. He joined the Lab as a Research Fellow in 2006 and worked there alongside Bill Hillier, particularly in his role as Programme Director, MSc Advanced Architectural Studies (now MSc Space Syntax), 2009–2014. His PhD was supervised by Alan Penn and Laura Vaughan, placing him in the third or fourth generation of space syntax scholars.

Sean Hanna is Professor of Design Computing at the Bartlett School of Architecture, University College London, having joined as a member of Bill Hillier's Space Syntax Laboratory in 2005. His research is primarily in developing computational methods for dealing with complexity in design and the built environment, including the comparative modelling of space, the use of machine learning and optimisation techniques for design and fabrication, and the analysis of human movement.

Frederico de Holanda has retired (2014) as Full Professor of Architecture at the School of Architecture, University of Brasília, from which he was also granted the title of Professor Emeritus (2019). He has published five books, among them the bi-lingual (Portuguese/English) *Oscar Niemeyer: Of glass and concrete*. He is currently Senior Research Fellow (topmost rank) of the Scientific-Technological Development National Council, Brazil. His MSc thesis (1977) and his PhD dissertation (1997) were both supervised by Bill Hillier.

Kayvan Karimi is a Professor of Urban Analytics and Spatial Design at the Bartlett School of Architecture, UCL, and a director of Space Syntax Limited, a spin-out company founded by Bill Hillier. Kayvan worked with Bill Hillier for over 20 years, initially at the company, and continued this collaboration after he joined UCL in 2011. Since 2017, Kayvan has served as Programme Director, MSc/MRes Space Syntax: Architecture and Cities MSc/MRes programmes, originally established by Bill Hillier in 1976.

Adrian Leaman is co-author with Bill Hillier of 'How is design possible?', 'Knowledge and design', 'The man–environment paradigm and its paradoxes', 'Architecture as a discipline', 'The architecture of architecture' and 'Space syntax'. These were all written when both Adrian and Bill worked at the RIBA in the early 1970s. Adrian's main interest then was spatial analysis. Subsequently, he has worked for Building Use Studies and the Usable Buildings Trust in promoting post-occupancy evaluation and user studies of buildings. He has over 150 publications.

Ann Legeby is Professor of Applied Urban Design at the School of Architecture, KTH Royal Institute of Technology, engaged in research and teaching. The research concerns society-space relations, focusing on the role of urban form in relation to social segregation and the conditions for everyday life. She has practised at Sweco Architects since 1998. Since 2018, she has led the research programme Applied Urban Design at KTH.

Lars Marcus is an architect and has been Professor in Urban Design at Chalmers University of Technology in Gothenburg since 2014. Earlier, he held the same position at the Royal Institute of Technology (KTH) in Stockholm. He is founder (2000) and partner (2005–2022) of the consultancy firm Spacescape, and also a fellow of the Royal Swedish Academy of Engineering Sciences.

Daniel R. Montello has been Professor of Geography, and Affiliated Professor of Psychological and Brain Sciences, at the University of California, Santa Barbara, since 1992. He was postdoctoral scholar at the Institute of Child Development, University of Minnesota, 1988–1991. He earned his doctorate in environmental psychology at Arizona State University in 1988, and his bachelor's degree in psychology from The Johns Hopkins University in 1981.

Vinicius Netto is a Principal Researcher at the Research Centre for Territory, Transport and Environment (CITTA | FEUP) at the University of Porto, Portugal. He is also an Associate Professor and former Director of the Graduate Programme in Architecture and Urbanism at Fluminense Federal University (UFF), Rio de Janeiro, Brazil. He earned his PhD in the early 2000s under the supervision of Bill Hillier at the Bartlett School of Architecture, University College London.

Alan Penn is Professor of Architectural and Urban Computing at the Bartlett School of Architecture, UCL, and a Director of Space Syntax Limited. After taking the MSc in Advanced Architectural Studies with Bill Hillier in 1981, he joined the Unit for Architectural Studies as Research Fellow in 1983 and has been associated with the development of space syntax ever since.

John Peponis is Professor of Architecture at the Georgia Institute of Technology since 1989. He served as a part-time member of the faculty at the National Technical University of Athens, 1992–2005. As a researcher and lecturer at the Bartlett (1978–1988), he was among the co-creators of space syntax and the first doctoral graduate supervised by Bill Hillier (1983).

Mahbub Rashid is a tenured full professor at the University of Kansas, Lawrence, KS, and a registered architect in the USA. He has served as the Dean of the School of Architecture and Design at the university since 2019. At the Georgia Institute of Technology, John Peponis served as his PhD supervisor and Bill Hillier was one of his external PhD examiners, along with Lionel March.

Ashraf M. Salama is Chair in Architecture and Urbanism and Head of Architecture and Built Environment at the University of Northumbria, Newcastle, UK. He was Head of the School of Architecture (2014–2020) and Director of Research (2020–2022) at the University of Strathclyde, Scotland. He is the chief editor of *Archnet-IJAR: International Journal of Architectural Research*. He is the UIA 2017 Recipient of Jean Tschumi Prize for Excellence in Architectural Education and Criticism.

Philip Steadman is Emeritus Professor of Urban and Built Form Studies at University College London, a position which he owes to the influence of Bill Hillier. He and Hillier never formally collaborated, but they debated and argued for three decades.

Laura Vaughan is Director of the Space Syntax Laboratory at the Bartlett School of Architecture, University College London, where she is Professor of Urban Form and Society. Following an architectural design degree at the Bezalel Academy of Art and Design, Jerusalem, Israel, she studied for an MSc and PhD at the Lab. After several years working with Bill Hillier at Space Syntax Limited, she returned to UCL in 2001 as lecturer and Programme Director, MSc Advanced Architectural Studies (now MSc/MRes Space Syntax). She has been the Lab's Director since 2014.

Wilfried Wang is co-founder of Hoidn Wang Partner, Berlin; Professor at Tongji University, Shanghai; President of CICA, International Committee of Architectural Critics; and a member of the Kungliga Akademien för de fria Konsterna, Stockholm, and the Akademie der Künste, Berlin. He holds an honorary doctorate from the Kungliga Tekniska Högskolan, Stockholm, and is an honorary member of the Portuguese Chamber of Architects. From 1980–1981 he was a student of Bill Hillier, Julienne Hanson and John Peponis.

List of published works by Bill Hillier

Chiaradia, A., Hillier, B., Schwander, C. and Wedderburn, M. 2009. 'Spatial centrality, economic vitality/viability: Compositional and spatial effects in Greater London'. In: Koch, D., Marcus, L. and Steen, J. (eds), *Seventh International Space Syntax Symposium*. Royal Stockholm: Institute of Technology.

Chiaradia, A., Hillier, B., Schwander, C. and Wedderburn, M. 2012. 'Compositional and urban form effects on centres in Greater London'. *Proceedings of the Institution of Civil Engineers: Urban Design and Planning*, 165, 21–42.

Croxford, B., Penn, A. and Hillier, B. 1996. 'Spatial distribution of urban pollution: Civilizing urban traffic'. *Science of The Total Environment*, 189, 3–9.

Friedrich, E., Hillier, B. and Chiaradia, A. 2009. 'Anti-social behaviour and urban configuration: Using space syntax to understand spatial patterns of socio-environmental disorder'. In: Koch, D., Marcus, L. and Steen, J. (eds), *Seventh International Space Syntax Symposium*. Stockholm: Royal Institute of Technology, 034:1–034:16.

Hanson, J. and Hillier, B. 1982. 'Domestic space organisation: Two contemporary space-codes compared'. *Architecture et Comportement/Architecture and Behaviour*, 2, 5–25.

Hanson, J. and Hillier, B. 1987. 'The architecture of community: Some new proposals on the social consequences of architectural and planning decisions'. *Architecture et Comportement/Architecture and Behaviour*, 3, 251–273.

Hanson, J. and Hillier, B. 1993. 'City of London: Shape and role'. *Urban Design Quarterly*, 48, 8–13.

Hillier, B. 1970. 'Psychology and the subject matter of architectural research'. In: Canter, D. (ed.), *Architectural psychology: Proceedings of the conference held at Dalandhui, University of Strathclyde*. University of Strathclyde and Royal Institute of British Architects, 25–29.

Hillier, B. 1971. '19th century building'. *Architectural Research and Teaching*, 1, 53–54.

Hillier, B. 1971. 'Architectural roles and profiles'. *Architectural Research and Teaching*, 1, 51–52.

Hillier, B. 1972. 'Architecture and engineering in environmental education'. *Architectural Research and Teaching*, 2, 111–118.

Hillier, B. 1973. 'In defence of space'. *RIBA Journal,* 11 (November), 539–544.

Hillier, B. 1985. 'The nature of the artificial'. *Geoforum*, 16, 163–178.

Hillier, B. 1985. 'Quite unlike the pleasures of scratching: Theory and meaning of architectural form'. *9H*, 7, 66–72.

Hillier, B. 1986. 'City of Alice's dreams'. *The Architects' Journal*, 184, 39–41.

Hillier, B. 1988. 'La morfologia urbana e le legge dell'oggetto'. In: Zanella, P. (ed.), *Morfologia dello spazio urbano.* Milan: Franco Angeli, 29–64.

Hillier, B. 1988. 'Against enclosure'. In: Teymur, N., Markus, T. and Wooley, T. (eds), *Rehumanizing housing.* London: Butterworth, 63–88.

Hillier, B. 1989. 'The architecture of the urban object'. *Ekistics*, 56, 5–21.

Hillier, B. 1992. 'Look back to London'. *The Architects' Journal*, 195, 42–46.

Hillier, B. 1993. 'Specifically architectural theory: A partial account of the ascent from building as cultural transmission to architecture as theoretical concretion'. *Harvard Architecture Review*, 9, 8–27.

Hillier, B. 1996. 'Cities as movement economies'. *URBAN DESIGN International*, 1, 49–60.

Hillier, B. 1996. *Space is the machine: A configurational theory of architecture*. Cambridge: Cambridge University Press.

Hillier, B. 1998. 'A note on the intuiting of form: Three issues in the theory of design'. *Environment and Planning B: Planning and Design*, 25(7), 37–40.
Hillier, B. 1998. 'From research to design: Re-engineering the space of Trafalgar Square'. *Urban Design Quarterly*, 68, 35–37.
Hillier, B. 1998. 'The hidden geometry of deformed grids: Or, why space syntax works, when it looks as though it shouldn't'. *Environment and Planning B: Planning and Design*, 26, 169–191.
Hillier, B. 1998. 'The reasoning art: Or, the need for an analytic theory of architecture'. In: Woyseth, H. and Noschis, K. (eds), *Architecture and teaching: Epistemological foundations*, EAAE Transactions on Architectural Education, no. 2. Lausanne: Comportements, 71–82.
Hillier, B. 1999. 'The need for domain theories'. *Environment and Planning B: Planning and Design*, 26, 163–167.
Hillier, B. 1999. 'The common language of space: A way of looking at the social, economic and environmental functioning of cities on a common basis'. *Journal of Environmental Sciences*, 11, 344–349.
Hillier, B. 1999. 'Centrality as a process: Accounting for attraction inequalities in deformed grids'. *URBAN DESIGN International*, 4, 107–127.
Hillier, B. 1999. 'Space as a paradigm for describing emergent structure in strongly relational systems'. In: Holanda, F. de, Amorim, L. and Dufaux, F. (eds), *Second International Space Syntax Symposium*. Brasília: University of Brasília, 1–16.
Hillier, B. 2000. 'Sicherheir durch Stadtdesign'. *Der Architekt*, 4, 41–43.
Hillier, B. 2001. 'Is there a "social device": A reflection on the work of Basil Bernstein'. In: Power, S., Aggleton, P., Brannen, J., Brown, A., Chisholm, L. and Mace, J. (eds), *A tribute to Basil Bernstein, 1924–2000*. London: Institute of Education, 54–59.
Hillier, B. 2001. 'A theory of the city as object: Or, how spatial laws mediate the social construction of urban space'. In: Peponis, J., Wineman, J. and Bafna, S. (eds), *Third International Space Syntax Symposium*. Atlanta, GA: Georgia Institute of Technology, 2.01–2.28.
Hillier, B. 2002. 'A theory of the city as object: Or, how spatial laws mediate the social construction of urban space'. *URBAN DESIGN International*, 3–4, 153–179.
Hillier, B. 2003. 'The knowledge that shapes the city: the human city beneath the social city'. *Fourth International Space Syntax Symposium*. London: University College London.
Hillier, B. 2003. 'The architectures of seeing and going: Or, are cities shaped by bodies or minds? And is there a syntax of spatial cognition?' In: Hanson, J. (ed.), *Fourth International Space Syntax Symposium*. London: University College London, 06:01–06:34.
Hillier, B. 2004. 'Can streets be made safe?' *URBAN DESIGN International*, 9, 31–45.
Hillier, B. 2004. 'What it looks like from here: Questions and conjectures about spatial cognition from a "syntactic" point of view'. *Keynote address to the Workshop on Mental Conceptions of Emergent Spaces in Architectural Design*. 18 July, MIT.
Hillier, B. 2005. 'Between social physics and phenomenology: Explorations towards an urban synthesis?' In: Van Nes, A. (ed.), *Fifth International Space Syntax Symposium*. Amsterdam: Techne Press, 3–24.
Hillier, B. 2005. 'The art of place and the science of space'. *World Architecture*, 185, 24–34 [in Chinese], 96–102 [in English].
Hillier, B. 2006. 'The golden age for cities? How we design cities is how we understand them'. *URBAN DESIGN*, 100 (Autumn), 16–19.
Hillier, B. 2007. 'What do we have to add to a social network to get a society? Answer: something like what we have to add to a spatial network to get a city'. *Conference on Spatial Information Theory*, COSIT'07. Melbourne, Australia, September 19–23.
Hillier, B. 2008. 'Space and spatiality: What the built environment needs from social theory'. *Building Research & Information*, 36, 216–230.
Hillier, B. 2008. 'New science of space and the art of place: Toward a space-led paradigm for researching and designing the city'. In: Haas, T. (ed.), *New urbanism and beyond: Designing cities for the future.* New York: Rizzoli, 30–39.
Hillier, B. 2009. 'What do we need to add to a social network to get a society? Answer: something like what you have to add to a spatial network to get a city'. In: Koch, D., Marcus, L. and Steen, J. (eds), *Seventh International Space Syntax Symposium*. Stockholm: Royal Institute of Technology, 042:1–042:15.
Hillier, B. 2009. 'The city as a socio-technical system: A spatial reformulation in the light of the levels problem and the parallel problem'. In: Stewart Hornsby, K., Claramunt, C., Denis, M. and Ligozat, G. (eds), *Conference on Spatial Information Theory (COSIT) 9th International Conference.* Aber Wrac'h, France, September 21–25.
Hillier, B. 2009. 'Spatial sustainability in cities: Organic patterns and sustainable forms'. In: Koch, D., Marcus, L. and Steen, J. (eds), *Seventh International Space Syntax Symposium*. Stockholm: Royal Institute of Technology, K01:1–K01:20.
Hillier, B. 2010. 'What do we need to add to a social network to get a society? Answer: something like what we have to add to a spatial network to get a city'. *Journal of Space Syntax*, 1, 41–58.
Hillier, B. 2011. 'Is architectural form meaningless: A configurational theory of generic meaning in architecture, and its limits'. *Journal of Space Syntax*, 2, 125–153.
Hillier, B. 2012. 'The city as a socio-technical system: A spatial reformulation in the light of the levels problem and the parallel problem'. In: Müller Arisona, S., Aschwanden, G., Halatsch, J. and Wonka, P. (eds), *Digital urban modelling and simulation.* Heidelberg, Dordrecht, London and New York: Springer, 24–48.
Hillier, B. 2012. 'The genetic code for cities: Is it simpler than we think?' In: Portugali, J., Meyer, H., Stolk, E. and Tan, E. (eds), *Complexity theories of cities have come of age: An overview with implications to urban planning and design.* Berlin and Heidelberg: Springer, 129–152.
Hillier, B. 2012. 'Studying cities to learn about minds: Some possible implications of space syntax for spatial cognition'. *Environment and Planning B: Planning and Design*, 39, 12–32.
Hillier, B. 2013. 'Credible mechanisms or spatial determinism'. *Cities*, 34, 75–77.
Hillier, B. 2013. 'Showing space: Or, can there be sciences of the non-discursive?' In: Richard, H., Elisabeth, N., Wolfram, P. and David, W. (eds), *Image and Imaging in Philosophy, Science and the Arts*, vol. 2, *Publications of the Austrian Ludwig Wittgenstein Society*, new ser., 17. Bergen: De Gruyter, 249–273.
Hillier, B. 2013. 'The now and future of space syntax: From structures and models to theory' (Keynote abstract). In: Young Ook Kim, Hoon Tae Park and Seo, K. W. (eds), *Ninth International Space Syntax Symposium*. South Korea: Sejong University Press, 00A1:2.
Hillier, B. 2014. 'Spatial analysis and cultural information: The need for theory as well as method in space syntax analysis'. In: Paliou, E., Lieberwirth, U. and Polla, S. (eds), *Spatial analysis and social spaces: Interdisciplinary approaches to the interpretation of prehistoric and historic built environments.* Berlin: Freie Universität Berlin, 19–47.
Hillier, B. 2016. 'What are cities for? And how does it relate to their spatial form?' *The Journal of Space Syntax*, 6, 199–212.
Hillier, B. 2016. 'The fourth sustainability, creativity: Statistical associations and credible mechanisms'. In: Portugali, J. and Stolk, E. (eds), *Complexity, cognition, urban planning and design.* Berlin and Heidelberg: Springer, 75–92.
Hillier, B., Burdett, R., Peponis, J. and Penn, A. 1987. 'Creating life: Or, does architecture determine anything?' *Architecture et Comportement/Architecture and Behaviour*, 3, 233–250.
Hillier, B., Greene, M. and Desyllas, J. 1999. 'Self-generated neighbourhoods: The role of urban form in the consolidation of informal settlements'. In: Hollanda, F., Amorim, L. and Dufaux, F. (eds), *Second International Space Syntax Symposium*. Brasília: University of Brasília, 1.1–24.
Hillier, B., Greene, M. and Desyllas, J. 2000. 'Self-generated neighbourhood consolidation in informal settlements'. *URBAN DESIGN International*, 5, 61–96.
Hillier, B. and Hanson, J. 1982. 'Space after modernism'. *9H (3)*, 15–20.
Hillier, B. and Hanson, J. 1984. *The social logic of space*. Cambridge, Cambridge University Press.
Hillier, B. and Hanson, J. 1987. 'A second paradigm'. *Architecture et Comportement/Architecture and Behaviour*, 3, 205–216.
Hillier, B. and Hanson, J. 1997. 'The reasoning art: Or, the need for an analytical theory of architecture'. In: Major, M. D., Amorim, L. and Dufaux, F. (eds), *First International Space Syntax Symposium*. London: University College London, 01.1–01.6.
Hillier, B. and Hanson, J. 1998. 'Space syntax as a research programme'. *Urban Morphology*, 2, 108–110.
Hillier, B., Hanson, J. and Graham, H. 1987. 'Ideas are in things: An application of the space syntax method to discovering house genotypes'. *Environment and Planning B: Planning and Design*, 14, 363–385.
Hillier, B., Hanson, J. and Peponis, J. 1984. 'What do we mean by building function?' In: Powell, J., Cooper, I. and Lera, S. (eds), *Designing for building utilisation.* London: Spon, 61–98.
Hillier, B., Hanson, J. and Peponis, J. 1987. 'Syntactic analysis of settlements'. *Architecture et Comportement/Architecture and Behaviour*, 3, 217–231.
Hillier, B., Hanson, J., Peponis, J., Hudson, J. and Burdett, R. 1983. 'Space syntax: A different urban perspective'. *The Architects' Journal*, 178(48), 47–54, 59–63.
Hillier, B. and Iida, S. 2005. 'Network and psychological effects in urban movement'. In: Van Nes, A. (ed.), *Fifth International Space Syntax Symposium*, Delft, Holland. Amsterdam: Techne Press, 553–564.
Hillier, B. and Iida, S. 2005. 'Network and psychological effects in urban movement'. In: Cohn, A. G. and Mark, D. M. (eds), *Spatial information theory*, *Lecture notes in computer science*, no. 3693. Berlin and Heidelberg: Springer, 475–490.
Hillier, B. and Jones, L. 1977. 'Architecture at the crossroad'. *New Scientist*, 74, 390–392.
Hillier, B. and Leaman, A. 1972. 'A new approach to architectural research'. *RIBA Journal*, 79 (12), 517–521.
Hillier, B. and Leaman, A. 1973. 'The man–environment paradigm and its paradoxes'. *Architectural Design*, 43, 507–511.

Hillier, B. and Leaman, A. 1973. ‘System, structure and transformation’. *Transactions of the Bartlett Society*, 9, 36–77.
Hillier, B. and Leaman, A. 1974. ‘How is design possible?’ *Journal of Architectural Research*, 3, 4–11.
Hillier, B. and Leaman, A. 1975. ‘The architecture of architecture: Foundations of a mathematical theory of artificial space’. In: Hawkes, D. (ed.), *Models and systems in architecture and building.* Hornby: Construction Press, 5–28.
Hillier, B. and Leaman, A. 1976. ‘Architecture as a discipline’. *Journal of Architectural Research*, 5, 28–32.
Hillier, B., Leaman, A., Stansall, P. and Bedford, M. 1976. ‘Space syntax’. *Environment and Planning B: Planning and Design*, 3, 147–185.
Hillier, B., Leaman, A., Stansall, P. and Bedford, M. 1978. ‘Reply to Professor Leach’. In: Green, D., Haselgrove, C. and Spriggs, M. (eds), *Social organisation and settlement: Contributions from anthropology, archaeology and geography.* Oxford: British Archaeological Reports, 403–405.
Hillier, B., Musgrove, J. and O’Sullivan, P. 1972. ‘Knowledge and design’. In: Mitchell, W. (ed.), *AIA-Architect-Researchers’ Conference. Environmental Design: Research and Practice. Proceedings of the Third Conference of the Environmental Design Research Association*. Los Angeles: University of California Press, 245–264.
Hillier, B., Musgrove, J. and O’Sullivan, P. 1984. ‘Knowledge and design’. In: Cross, N. (ed.), *Developments in design methodology*. New York: John Wiley & Sons, 245–264.
Hillier, B. and Netto, V. M. 2001. ‘Society seen through the prism of space: Outline of a theory of society and space’. In: Peponis, J., Wineman, J. and Bafna, S. (eds), *Third International Space Syntax Symposium*. Atlanta, GA: Georgia Institute of Technology, 13.1–13.27.
Hillier, B. and Netto, V. 2002. ‘Society seen through the prism of space: Outline of a theory of society and space’. *URBAN DESIGN International*, 7, 181–203.
Hillier, B. and Penn, A. 1991. ‘Visible colleges: Structure and randomness in the place of discovery’. *Science in Context*, 4, 23–49.
Hillier, B. and Penn, A. 1992. ‘Dense civilisations: Or, the shape of cities in the 21st century’. *Applied Energy*, 43, 41–66.
Hillier, B. and Penn, A. 1994. ‘Virtuous circles, building sciences and the science of buildings: Using computers to integrate product and process in the built environment’. *Design Studies*, 15, 332–365.
Hillier, B. and Penn, A. 2004. ‘Rejoinder to Carlo Ratti’. *Environment and Planning B: Planning and Design*, 31, 501–511.
Hillier, B. and Penn, A. 2010. ‘Exemplars in design’. In: Brooks, F. P. (ed.), *The design of design: Essays from a computer scientist.* Upper Saddle River, NJ: Addison-Wesley, 153–166.
Hillier, B., Penn, A., Dalton, N., Chapman, D. and Redfern, F. 1995. ‘Graphical knowledge interfaces: The extensive and intensive use of precedent databases in architecture and urban design’. In: Koutamanis, A., Timmermans, H. and Vermeulen, I. (eds), *Visual databases in architecture: Recent advances in design and decision making.* Aldershot: Avebury, 197–227.
Hillier, B., Penn, A., Grajewski, T. and Xu, J. 1991. *Brindleyplace, Birmingham*. London: Unit for Architectural Studies, UCL.
Hillier, B., Penn, A., Hanson, J., Grajewski, T. and Xu, J. 1993. ‘Natural movement: Or, configuration and attraction in urban pedestrian movement’. *Environment and Planning B: Planning and Design*, 20, 29–66.
Hillier, B., Peponis, J. and Simpson, J. 1982. ‘National gallery schemes analysed’. *The Architects’ Journal*, 176, 38–40.
Hillier, B. and Sahbaz, O. ‘High resolution analysis of crime patterns in urban street networks: An initial statistical sketch from an ongoing study of a London borough’. In: Van Nes, A. (ed.), *Fifth International Space Syntax Symposium*, 2005 Delft, Holland. Amsterdam: Techne Press, 451–478.
Hillier, B. and Sahbaz, O. 2009. ‘Crime and urban design: An evidence-based approach’. In: Cooper, R., Evans, G. and Boyko, C. (eds), *Designing sustainable cities: Decision-making tools and resources for design.* Chichester: Wiley Blackwell, 163–186.
Hillier, B. and Sahbaz, O. 2012. ‘Safety in numbers: High resolution analysis of crime patterns in urban street networks’. In: Ceccato, V. (ed.), *Urban fabric of crime and fear.* Berlin and Heidelberg: Springer, 451–478.
Hillier, B. and Shu, S. 1999. ‘Designing for secure spaces’. *Planning in London*, 29 (April), 36–38.
Hillier, B. and Shu, S. 2000. ‘Crime and urban layout: The need for evidence’. In: Maclaren, V., Ballantyne, S. and Pease, K. (eds), *Secure foundations: Key issues in crime prevention and community safety.* London: Institute of Public Policy Research, 224–248.
Hillier, B. and Stonor, T. 2010. ‘Space syntax: Strategic urban design’. *City Planning Review*, 59, 7–11.
Hillier, B. and Stutz, C. 2005. ‘New methods in space syntax’. *Urban Design*, 93, 32–33.
Hillier, B., Turner, A., Yang, T. and Tae-Park, H. 2007. ‘Metric and topo-geometric properties of urban street networks: Some convergences, divergences, and new results’. In: Kubat, A. S. (ed.), *Sixth International Space Syntax Symposium*. Istanbul: Technical University of Istanbul, 001:1–001:22.
Hillier, B., Turner, A., Yang, T. and Tae-Park, H. 2010. ‘Metric and topo-geometric properties of urban street networks: Some convergences, divergences and new results’. *Journal of Space Syntax*, 1, 258–279.
Hillier, B. and Tzortzi, K. 2006. ‘Spatial syntax: The language of museum space’. In: Macdonald, S. (ed.), *A companion to museum studies.* Oxford: Blackwell Publishing, ch 17.
Hillier, B. and Tzortzi, K. 2016. ‘From exhibits to spatial culture: An exploration of performing arts collections in museums’. *Journal of Space Syntax*, 7, 71–86.
Hillier, B. and Vaughan, L. 2007. ‘The city as one thing’. In: Vaughan, L. (ed.), ‘The syntax of segregation’. Special issue, *Progress in Planning*, 67, 205–230.
Hillier, B., Young, T. and Turner, A. 2012. ‘Normalising least angle choice in DepthMap – and how it opens up new perspectives on the global and local analysis of city space’. *Journal of Space Syntax*, 3(2), 155–193.
Major, M. D., Penn, A. and Hillier, B. 1999. ‘The urban village and the city of tomorrow revisited’. In: Holland, F., Amorim, L. and Dufaux, F. (eds), *Second International Space Syntax Symposium*. Brasília: University of Brasília, 1–18.
O’Sullivan, P., Territt, C., Musgrove, J., Hillier, B. and Leaman, A. 1975. ‘Architectural research: Problems of organization and funding in the United Kingdom’. *Journal of Architectural Research*, 4, 41–46.
Orhun, D., Hillier, B. and Hanson, J. 1995. ‘Spatial types in traditional Turkish houses’. *Environment and Planning B: Planning and Design*, 22, 475–498.
Orhun, D., Hillier, B. and Hanson, J. 1996. ‘Socialising spatial types in traditional Turkish houses’. *Environment and Planning. B, Planning and Design.* 23, 329–351.
Ortiz-Chao, C. and Hillier, B. 2007. ‘In search of patterns of land-use in Mexico City using logistic regression at the plot level’. In: Kubat, A. S. (ed.), *Sixth International Space Syntax Symposium*. Istanbul: Technical University of Istanbul, 016:01–016:16.
Penn, A. and Hillier, B. 1992. ‘The social potential of buildings: Spatial structure and the innovative millieu in scientific research laboratories’. *Corporate Space and Architecture: International Symposium*. Ministère de l’Equipement du Logement et des Transports, 39–43.
Penn, A., Hillier, B., Banister, D. and Xu, J. 1998. ‘Configurational modelling of urban movement networks’. *Environment and Planning B: Planning and Design*, 24, 59–84.
Raford, N. and Hillier, B. 2010. ‘Description and discovery in socio-spatial analysis: Space syntax’. In: Walford, G., Tucker, E. and Viswanathan, M. (eds), *The SAGE handbook of measurement.* London: Sage Publications, 265–282.
Rodriguez, C., Sakr, F. L., Griffiths, S. and Hillier, B. 2012. ‘The relationship of spatial configuration and socio-economic conditions in São Paulo, Brazil’. In: Greene, M., Reyes, J. and Castro, A. (eds), *Eighth International Space Syntax Symposium*, 3–6 January. Chile: Pontificia Universidad Católica de Chile.
Serra, M., Hillier, B. and Karimi, K. 2015. ‘Exploring countrywide spatial systems: Spatio-structural correlates at the regional and national scales’. In: Karimi, K., Vaughan, L., Sailer, K., Palailogou, G. and Bolton, T. (eds), *Proceedings of the Tenth International Space Syntax Symposium*. London: University College London, 84:1–84:18.
Steadman, P. and Hillier, B. 2002. ‘Research assessment under the microscope: Disturbing findings and distorting effects: The 2001 RAE dissected: some facts and figures’. *Architectural Research Quarterly*, 6, 203–207.
Stillman, J., Hillier, B., Jones, L., Bessey, G., Nelson, J., Kadleigh, S., Kimber, K., Hobbs, C., Wigglesworth, G. and Wise, D. 1977. ‘New look at building technology’. *RIBA Journal*, 84, 305–307.
Turner, A., Penn, A. and Hillier, B. 2005. ‘An algorithmic definition of the axial map’. *Environment and Planning B: Planning and Design*, 32, 425–444.
Varoudis, T., Law, S., Karimi, K., Hillier, B. and Penn, A. 2013. ‘Space syntax angular betweenness centrality revisited’. In: Young Ook Kim, Hoon Tae Park and Seo, K. W. (eds), *Ninth International Space Syntax Symposium*. South Korea: Sejong University Press, 00A1:2.
Yang, T. and Hillier, B. 2007. ‘The fuzzy boundary: The spatial definition of urban areas’. In: Kubat, A. S. (ed.), *Sixth International Space Syntax Symposium*. Istanbul: Technical University of Istanbul, 091.1–091.16.
Yang, T. and Hillier, B. 2015. ‘Spatial discontinuity in the multiscaled area structures of the central historic districts of London and Beijing’. In: Jin Duan, Xiaoling Dai, Jingjing Hu, Ning Liu, Runqing Shao, Qiang Sheng, Haofeng Wang, Sinan Yuan and Yu Zhuang (eds), *Twelfth International Space Syntax Symposium*. Beijing: Sejong University Press, 2429–2458.

Space Syntax index

References to figures are in *italics*; references to tables are in **bold**; references to notes are indicated by n.

Professor Bill Hillier spent most of his career at The Bartlett, University College London, where he founded and developed, with a team of colleagues, an original research programme that set the study of architecture on a firm scientific basis. His transformational way of thinking about buildings and cities influenced generations of scholars, researchers and practitioners within the built environment disciplines and way beyond – in fields ranging from archaeology and biology to physics and zoology.

Space Syntax: Selected papers by Bill Hillier provides a canon of works that reflects the progression of Hillier's ideas from the early publications of the 1970s to his most recent work, published before his death in 2019. This selection of influential works ranges from his papers on architecture as a professional and research discipline, through to his later papers that present a theory of the spatial structure of the city and its social functions. By bringing together writing from across his career-span of half a century, with specially commissioned introductions by a wide range of international experts in the field, we are able to contextualise and show the evolution of Hillier's key ideas.

Laura Vaughan is Professor of Urban Form and Society and Director of the Space Syntax Lab, The Bartlett School of Architecture, UCL.

John Peponis is Professor of Architecture at the Georgia Institute of Technology.

Ruth Conroy Dalton is Professor of Architecture at the University of Northumbria at Newcastle.

Free open access version available from www.uclpress.co.uk

UCLPRESS

Cover image: Ruth Conroy Dalton

Cover design: www.hayesdesign.co.uk