

ΠΛΗΡΟΦΟΡΙΑ

Θεωρίες Πληροφορίας και Επιστήμη των ΗΥ

Ιστορία
Κωδικοποίηση
Φυσική
Μαθηματικά
Επιστημολογία
Επιστήμη των ΗΥ

Computation  Information



Προσεγγίσεις /Θεωρίες Πληροφορίας:

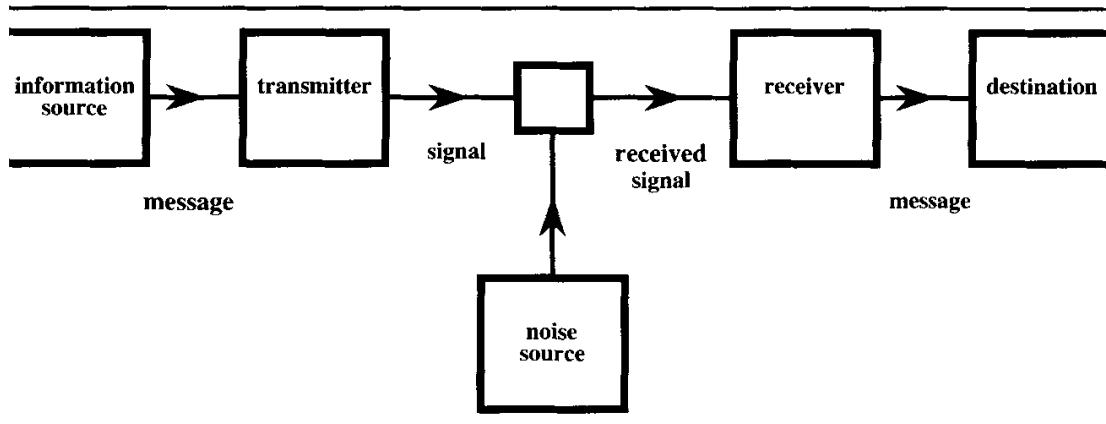
- Μαθηματική –ποσοτική Θεωρία της Επικοινωνίας Data / Signals
- Αλγοριθμική
- Πιθανοκρατική
- Modal
- Συστημική
- Inferential
- Σημασιολογική

ΠΟΣΟΤΙΚΗ ΘΕΩΡΙΑ

Shannon:

*«The word **'information'** has been given different meanings by various writers in the general field of information theory. It is likely that at least a number of these will prove sufficiently useful in certain applications to deserve further study and permanent recognition. It is hardly to be expected that a single concept of information would satisfactorily account for the numerous possible applications of this general field»*

*«Frequently the messages have meaning: that is they are referred to or correlated according to some system with certain physical or conceptual entities. These **semantic** aspects of communication **are irrelevant** to the **engineering** problem. The significant aspect is that the actual message is **one selected** from a set of possible messages»*



άπογοητευτική και παράξενη “disappointing & bizarre” (Weaver)

Μαθηματική Θεωρία επιτρέπει εξαγωγή Θεωρημάτων

Πληροφορία ως Αβεβαιότητα

Δυο βασικές μετρήσεις επιτρέπει η Ποσοτική Θεωρά Πληροφορίας:

- (1) Ποσότητα πληροφορίας που μπορεί να κωδικοποιηθεί ή ενσωματωθεί σε ένα σύστημα συμβόλων: bits
- (2) Βαθμοί Αβεβαιότητας: entropy

Ο Shannon όρισε την ποσότητα πληροφορίας σε ένα μήνυμα ως το αρνητικό (1-) του λογάριθμου (βάσης 2) της πιθανότητας του να συμβεί (το περιγραφόμενο συμβάν/κατάσταση) εκπεμπόμενο από ένα πομπό (πηγή) μέσω ενός μέσου (καναλιού) –μετρούμενο σε ‘bits’

ΣΣ: bits, nats, Hartleys κλπ φυσικοί λογαριθμοί (ως ισοδύναμες μετρήσεις)

Επίσης ΣΣ! Ενδιάμεσες καταστάσεις και Λογική πολλών τιμών

Η Πληροφορία ως ποσότητα που εμπεριέχεται σε (φέρεται από) μια σειρά στοιχείων μετράται λογαριθμικά και είναι συνάρτηση του μήκους της σειράς:

$\text{Info}(s) = \log_2 \text{Inv Prob}(s) = \text{Length}(s).$

$s = \text{string}$ (συμβολοσειρά 0-1, ναι-οχι, on-off, spin up-spin down, κλπ)

$\text{Info}(E) = \log_2[1/\text{Prob}(E)]$ μέτρηση με αντίστροφη πιθανότητα

$E = \text{event}$ (συμβαν/κατάσταση)

ΣΣ! Στην περίπτωση δυαδικών σειρών (1-0 κλπ) η ποσότητα πληροφορίας ισούται (είναι το) μήκος της σειράς π.χ. **101** = 3 (λογαριθμικά η πιθανότητα 1/8 της συγκεκριμένης σειράς)

a source emitting signals with certain frequencies, say a 'language' viewed as a global text producer, and the information which a receiver picks up from this is measured in terms of expected reduction of uncertainty

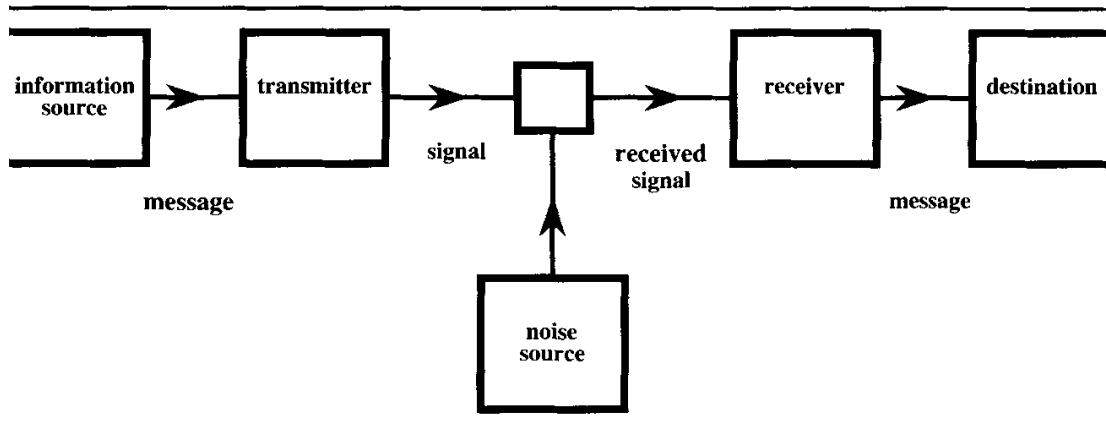
[information theory is the thermodynamics of code strings, while thermodynamics is the information theory of particles in space]

Suppose X is a set of messages x_i ($i = 1, \dots, n$) the communication entropy of X is

$$H(X) = - \sum_{i=1, n} P(x_i) \log P(x_i)$$

The Maximal entropy of a set of n messages, if $P(x_i) = 1/n$ for each i :

$$H_{\max}(X) = -n(1/n) \log (1/n) = \log n$$



A source **S**, which generates the message to be received at the destination.

A transmitter **T**, which turns the message generated at the source into a signal to be transmitted. When information is encoded, coding is also implemented by this system.

A channel **CH**, that is, the medium used to transmit the signal from the transmitter to the receiver.

A receiver **R**, which reconstructs the message from the signal.

A destination **D**, which receives the message

Η πηγή **S** είναι ένα σύστημα η καταστάσεων s_i , καθε μια των οποίων έχει την δική της πιθανότητα $p(s_i)$; Η ακολουθία καταστάσεων καλείται *μήνυμα(τα)*.

Η ποσότητα πληροφορίας που παράγεται από την **S** με τα συμβάντα s_i είναι $I(s_i) = -\log_2 p(s_i)$,

οπότε και η εντροπία της πηγής S ορίζεται ως μια μέση τιμή δηλ., το άθροισμα των επι μέρους ποσοτήτων πληροφορίας με συντελεστή βαρύτητας την αντίστοιχη πιθανότητα κάθε s_i :

$$H(S) = -\sum p(s_i)\log_2 p(s_i)$$

Και κατ'αναλογία ο προορισμός **D** ορίζεται με τον ίδιο τρόπο. Οπότε η εντροπία της απόληξης D ισούται με την μέση τιμή της πληροφορίας που παραλαμβάνεται στο σημείο **D**:

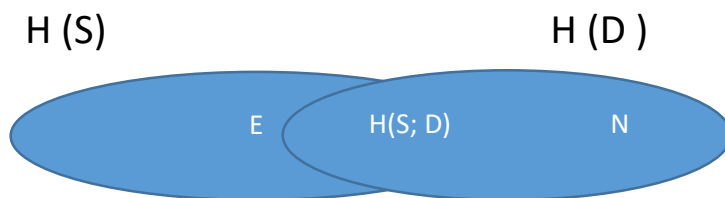
$$H(D) = -\sum p(d_j)\log_2 p(d_j)$$

$H(S;D)$ είναι η λεγόμενη mutual information: η μέση τιμή ποσότητας πληροφορίας που παράγεται στην S και παραλαμβάνεται στην D .

E = equivocation: μέση τιμή πληροφορίας που παράγεται από την πηγή αλλά δεν παραλαμβάνεται από την D .

N = noise: η μέση τιμή ποσότητας πληροφορίας που παραλαμβάνεται στην D και που δεν έχει παραχθεί στην πηγή S .

E και N μετρούν την σχέση εξάρτησης μεταξύ S και D :



$$H(S;D) = H(S) - E = H(D) - N$$

Όταν S και D είναι εντελώς ανεξάρτητα, οι τιμές E και N είναι μέγιστες: ($E=H(S)$ και $N=H(D)$), και η τιμή του $H(S;D)$ είναι στο ελάχιστο ($H(S;D)=0$).

Όταν η αλληλεξάρτηση μεταξύ S και D είναι η μέγιστη, οι τιμές E και N είναι ελάχιστες, ($E=N=0$), και η τιμή $H(S;D)$ είναι στο μέγιστο ($H(S;D)=H(S)=H(D)$).

ΑΛΓΟΡΙΘΜΙΚΗ ΘΕΩΡΙΑ της ΠΛΗΡΟΦΟΡΙΑ

Εμπνευστές: Kolmogorov, Chaitin, Solomonoff

The algorithmic information of a sequence is the length of the shortest algorithm that generates the sequence

“Κ-πολυπλοκότητα»

K C S: information in a string X = the length of the shortest code of a program that would lead a universal Turing Machine to output string X