

Δένδρα

Γιώργος Στάμου

Εργαστήριο Συστημάτων Τεχνητής Νοημοσύνης και Μάθησης
Τομέας Τεχνολογίας Πληροφορικής και Υπολογιστών
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Εθνικό Μετσόβιο Πολυτεχνείο



Εισαγωγή

- Τα **δένδρα** είναι δομές δεδομένων που αναπαριστούν κάποιου είδους ιεραρχία
- Είναι ιδιαίτερα χρήσιμα για μια μεγάλη κατηγορία υπολογισμών και προβλημάτων
- Μπορούμε να χρησιμοποιήσουμε τα δένδρα ως **αποτελεσματικές** δομές αναπαράστασης και διαχείρισης δεδομένων
 - Μπορούν να υποστηρίξουν αποδοτικά: την εισαγωγή, διαγραφή ή τροποποίηση στοιχείων **και** την αναζήτησή τους

Ορισμός δένδρου

- Ένα **δένδρο** T είναι μία πεπερασμένη, πιθανά κενή λίστα

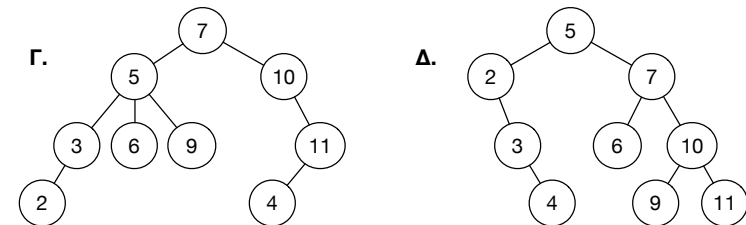
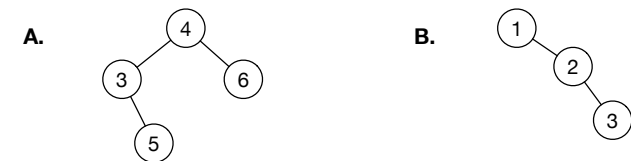
$$T = [r; T_1, T_2, \dots, T_n]$$

με τις εξής ιδιότητες:

- Το ορισμένο στοιχείο της λίστας (αν υπάρχει), δηλαδή το r , καλείται **ρίζα** του δένδρου
- Καθένα από τα υπόλοιπα στοιχεία $T_1, T_2, \dots, T_n$ της λίστας ($n \geq 0$) είναι δένδρα (τα ονομάζουμε **υποδένδρα** του T)

(Αναφέρουμε τα ορισμένα στοιχεία της λίστας (μαζί με αυτά όλων των υποδένδρων) ως **κόμβους** του δένδρου)

Παραδείγματα δένδρων



Παράδειγμα δενδρικής δομής

```
D: begin
  E: begin
    F: statement
  end;
  G: begin
    H: begin
      I: statement
    end;
    J: begin
      K: statement;
      L: statement
    end;
    M: statement
  end
end
end
```

Ορολογία

Έστω ένα δένδρο $T = [r; T_1, T_2, \dots, T_n], n \geq 0$.

- Ο **βαθμός** ενός κόμβου είναι ο αριθμός των υποδένδρων που συνδέονται με αυτόν τον κόμβο. Για παράδειγμα, ο βαθμός της ρίζας r του δένδρου T είναι n .
- Ένας κόμβος με βαθμό μηδέν δεν έχει κανένα υποδένδρο. Ένας τέτοιος κόμβος ονομάζεται **φύλλο**.
- Κάθε μία από τις ρίζες r_i των υποδένδρων T_i του δένδρου T ονομάζεται **παιδί** του r . (Ο όρος **εγγόνι** ορίζεται κατά παρόμοιο τρόπο.)

Ορολογία

Έστω ένα δένδρο $T = [r; T_1, T_2, \dots, T_n], n \geq 0$.

- Η ρίζα r του δένδρου T είναι ο **πατέρας** όλων των ριζών r_i των υποδένδρων $T_i, 1 \leq i \leq n$. (Ο όρος **παππούς** ορίζεται κατά παρόμοιο τρόπο.)
- Οι ρίζες r_i και r_j δύο διαφορετικών υποδένδρων T_i και T_j του δένδρου T ονομάζονται **αδέρφια**.

Μονοπάτι δένδρου

Έστω ένα δένδρο $T = [r; T_1, T_2, \dots, T_n], n \geq 0$ το οποίο περιέχει ένα σύνολο κόμβων R .

- **Μονοπάτι** του δένδρου είναι μία μη-κενή ακολουθία κόμβων:
 $P = r_1, r_2, \dots, r_k$
όπου r_i κόμβοι του δένδρου, για $1 \leq i \leq k$ έτσι ώστε ο i -οστός κόμβος στην ακολουθία, r_i , να είναι ο πατέρας του $(i+1)$ -οστού κόμβου στην ακολουθία r_{i+1} .
- Το **μήκος** του μονοπατιού P είναι $k - 1$.

Επιπλέον ορολογία

Έστω ένα δένδρο $T = [r; T_1, T_2, \dots, T_n]$, $n \geq 0$ το οποίο περιέχει ένα σύνολο κόμβων R .

- Το **επίπεδο** (level) ή **βάθος** (depth) ενός κόμβου r_i του δένδρου T είναι το μήκος του μοναδικού μονοπατιού από την ρίζα r στον κόμβο r_i . Για παράδειγμα, η ρίζα του δένδρου T είναι στο επίπεδο μηδέν και οι ρίζες των υποδένδρων του T είναι όλες στο επίπεδο ένα.
- Το **ύψος** (height) ενός κόμβου r_i του δένδρου T είναι το μήκος του μεγαλύτερου σε μήκος μονοπατιού από το κόμβο r_i σε κάποιο φύλλο. Επομένως, όλα τα φύλλα έχουν ύψος ίσο με μηδέν.

Επιπλέον ορολογία

Έστω ένα δένδρο $T = [r; T_1, T_2, \dots, T_n]$, $n \geq 0$ το οποίο περιέχει ένα σύνολο κόμβων R .

- Το **ύψος** ενός δένδρου T είναι το ύψος της ρίζας του r .
- Θεωρούμε (κατά σύμβαση) ότι το κενό δένδρο έχει ύψος -1.
- Έστω δύο κόμβοι r_i και r_j ενός δένδρου T .
 - Ο κόμβος r_i καλείται **πρόγονος** (ancestor) του κόμβου r_j εάν υπάρχει ένα μονοπάτι στο T από τον r_i στον r_j .
 - Ο κόμβος r_i καλείται **γνήσιος πρόγονος** (proper ancestor) εάν υπάρχει ένα μονοπάτι p στο δένδρο T από τον κόμβο r_i στον κόμβο r_j με μήκος μονοπατιού p να είναι διάφορο του μηδενός.

Επιπλέον ορολογία

Έστω ένα δένδρο $T = [r; T_1, T_2, \dots, T_n]$, $n \geq 0$ το οποίο περιέχει ένα σύνολο κόμβων R .

- Ο κόμβος r_j καλείται **απόγονος** (descendant) του κόμβου r_i εάν υπάρχει ένα μονοπάτι στο T από τον r_i στον r_j .
- Καθώς οι κόμβοι r_i και r_j μπορεί να συμπίπτουν, ένας κόμβος είναι απόγονος του εαυτού του.
- Ο κόμβος r_j καλείται **γνήσιος απόγονος** (proper descendant) εάν υπάρχει ένα μονοπάτι p στο T από τον r_i στον r_j τέτοιο ώστε το μήκος του μονοπατιού p να είναι διάφορο του μηδενός.

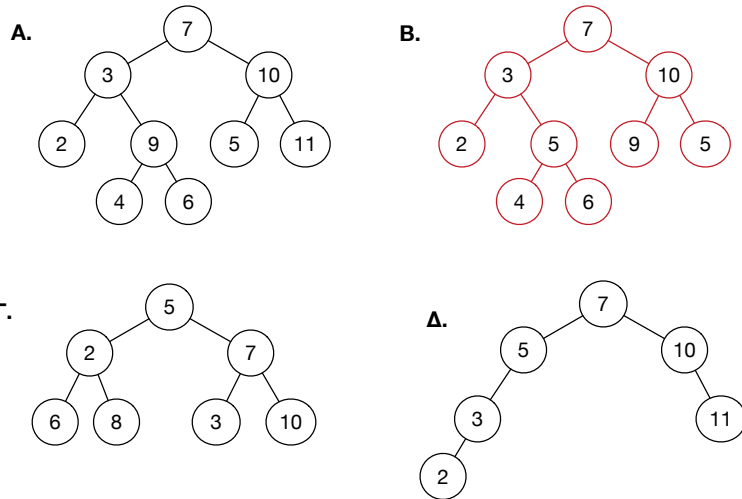
Διαδικά δένδρα

- Ένα δυαδικό δένδρο T είναι ένα δένδρο του οποίου κάθε κόμβος αποτελείται από μία ρίζα, r , και αποκλειστικά και μόνο από δύο επιπλέον δυαδικά δένδρα T_L και T_R δηλαδή

$$T = [r; T_L, T_R]$$

- Το δένδρο T_L ονομάζεται αριστερό υποδένδρο του T , και το δένδρο T_R ονομάζεται δεξιό υποδένδρο του T .
- Προφανώς, τα δυαδικά δένδρα είναι υποπερίπτωση των δένδρων $T = [r; T_1, T_2, \dots, T_n]$, $n \geq 0$ με $n=2$ παντού.

Δυαδικά δένδρα



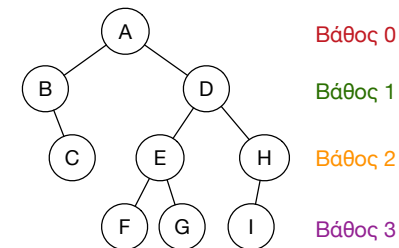
Ιδιότητες δυαδικών δένδρων

- Ένα δυαδικό δένδρο ύψους h έχει το πολύ $2^{h+1} - 1$ κόμβους.
- Αντιστρόφως, το ύψος ενός δυαδικού δένδρου με n κόμβους είναι τουλάχιστον $\lceil \log_2 n + 1 \rceil - 1$.
- Δηλαδή το ύψος ενός δυαδικού δένδρου το οποίο έχει n κόμβους είναι της τάξης $\Omega(\log n)$.
- Ένα δυαδικό δένδρο ύψους $h \geq 0$ έχει το πολύ 2^h φύλλα.
- Αντιστρόφως, το ύψος ενός δυαδικού δένδρου το οποίο έχει k φύλλα είναι τουλάχιστον $\lceil \log_2 k \rceil$.
- Έτσι, το ύψος ενός δυαδικού δένδρου με k φύλλα είναι της τάξης $\Omega(\log k)$.

Διάσχιση δένδρων

- Η διαδικασία κατά την οποία συστηματικά επισκεπτόμαστε όλους τους κόμβους του δένδρου ονομάζεται **διάσχιση δένδρου** (tree traversal).
- Δύο διαφορετικές μέθοδοι για να επισκεφθεί κανείς συστηματικά όλους τους κόμβους ενός δένδρου:
 - διάσχιση **κατά βάθος** (depth-first traversal)
 - διάσχιση **κατά πλάτος** (breadth-first traversal)
- Μέθοδοι κατά βάθος διάσχισης:
 - **προδιατεταγμένη** διάσχιση (preorder traversal)
 - **ενδοδιατεταγμένη** διάσχιση (inorder traversal)
 - **μεταδιατεταγμένη** διάσχιση (postorder traversal)

Διάσχιση δένδρων

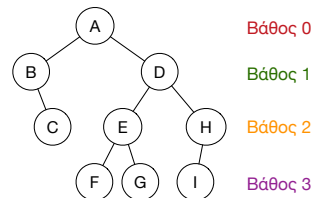


Διάσχιση κατά βάθος

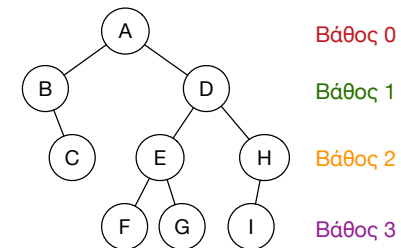
- Κατά την προδιατεταγμένη διάσχιση ενός γενικού δένδρου:
 - Επισκεπτόμαστε πρώτα τη ρίζα και στη συνέχεια
 - Εκτελούμε μία προδιατεταγμένη διάσχιση διαδοχικά σε κάθε ένα από τα υποδένδρα της ρίζας, με τη σειρά που δίνονται

- Στην περίπτωση ενός δυαδικού δένδρου:

- Επισκεπτόμαστε πρώτα τη ρίζα
- Διασχίζουμε το αριστερό υποδένδρο
- Διασχίζουμε το δεξί υποδένδρο



Διάσχιση δένδρων (προδιατεταγμένη)



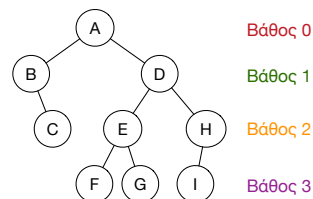
A, B, C, D, E, F, G, H, I.

Διάσχιση κατά βάθος

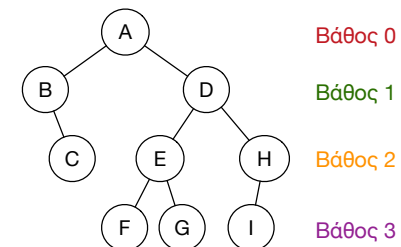
- Στη μεταδιατεταγμένη διάσχιση ενός γενικού δένδρου:
 - Εκτελούμε μία μεταδιατεταγμένη διάσχιση διαδοχικά σε κάθε ένα από τα υποδένδρα της ρίζας με τη σειρά που αυτά δίνονται
 - Επισκεπτόμαστε τη ρίζα

- Στην περίπτωση ενός δυαδικού δένδρου:

- Διασχίζουμε το αριστερό υποδένδρο
- Διασχίζουμε το δεξί υποδένδρο
- Επισκεπτόμαστε τη ρίζα



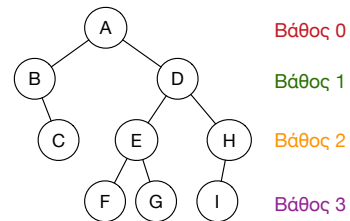
Διάσχιση δένδρων (μεταδιατεταγμένη)



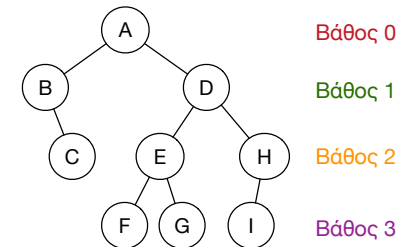
C, B, F, G, E, I, H, D, A.

Διάσχιση κατά βάθος

- Η ενδοδιατεταγμένη διάσχιση έχει νόημα μόνο στην περίπτωση ενός δυαδικού δένδρου:
 - Διασχίζουμε το αριστερό υποδένδρο
 - Επισκεπτόμαστε τη ρίζα
 - Διασχίζουμε το δεξί υποδένδρο.



Διάσχιση δένδρων

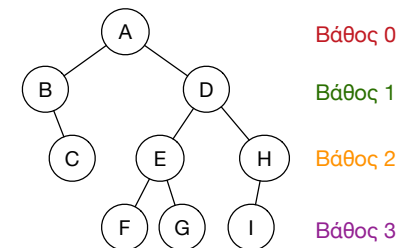


B, C, A, F, E, G, D, I, H.

Διάσχιση κατά πλάτος

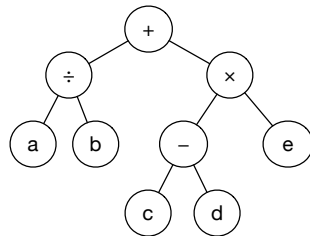
- Ενώ οι μέθοδοι διάσχισης κατά βάθος ορίζονται αναδρομικά, η *αναζήτηση κατά πλάτος* (breadth-first traversal) γίνεται καλύτερα κατανοητή χωρίς τη χρήση αναδρομής.
- Για να εκτελέσουμε τη διάσχιση ενός δένδρου κατά πλάτος, επισκεπτόμαστε τους κόμβους του δένδρου κατά σειρά βάθους.
- Η διάσχιση κατά πλάτος πρώτα επισκέπτεται όλους τους κόμβους που έχουν βάθος 0 (δηλαδή τη ρίζα), στη συνέχεια όλους τους κόμβους που έχουν βάθος 1 και ούτω καθεξής.
- Για κάθε τιμή βάθους επισκεπτόμαστε τους κόμβους από τα αριστερά προς τα δεξιά.

Διάσχιση δένδρων



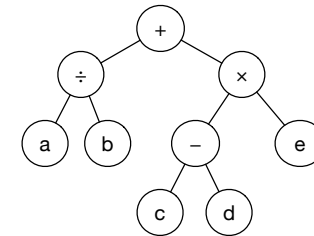
A, B, D, C, E, H, F, G, I.

Δένδρα εκφράσεων (προθεματική)



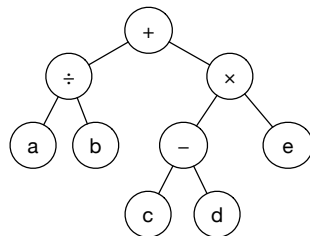
$+(\div(a, b), \times(- (c, d), e))$

Δένδρα εκφράσεων (ενδοθεματική)



$((a \div b) + ((c - d) \times e))$

Δένδρα εκφράσεων (μεταθεματική)



$ab \div cd - e \times +$