



# Περιεχόμενα

- Εισαγωγή
- Ορισμοί
- Προτάσεις
- Δυαδικά Δένδρα
  - Διάσχιση/Επίσκεψη
  - Αποθήκευση
  - Δυαδικά Δένδρα Αναζήτησης
- Μη Δυαδικά Δένδρα
  - Πολυκατευθυνόμενα Δένδρα (Δένδρα  $m$ -δρόμων)

## Εισαγωγή

Στην προηγούμενη ενότητα εξετάσαμε δομές οι οποίες ήταν **γραμμικές** (ο κάθε κόμβος ήταν ο ένας μετά τον άλλο). Σε αυτή την ενότητα θα ασχοληθούμε με μία **μη-γραμμική** δομή το **δένδρο** (tree). Η έννοια του δένδρου είναι συνδεδεμένη με την ιεραρχία.

## Ορισμοί

### ❖ Δέντρο (tree)

Ένα δένδρο  $T$  είναι ένα πεπερασμένο σύνολο από έναν ή περισσότερους κόμβους. Ένας ειδικός κόμβος  $t$  του δένδρου ονομάζεται **ρίζα (root)**, ενώ οι υπόλοιποι κόμβοι  $T - \{t\}$  μπορούν να διαμερισθούν σε ξένα μεταξύ τους υποσύνολα  $T_1, T_2, \dots, T_N$ , όπου κάθε υποσύνολο είναι με τη σειρά του ένα δένδρο. Τα δένδρα αυτά ονομάζονται **υποδένδρα (subtrees)** του αρχικού δένδρου.

### ❖ Βαθμός του Κόμβου (node degree)

Είναι ο αριθμός των υποδένδρων που αρχίζουν από τον κόμβο.

### ❖ Βαθμός Δένδρου (tree degree)

Είναι ο μέγιστος βαθμός από όλους τους βαθμούς των κόμβων.

Τα δυαδικά (binary trees) δένδρα έχουν βαθμό 2, τα τετραδικά (quadtrees) 4, τα οκταδικά 8 (octtrees).

## Ορισμοί

### ❖ Φύλλα (leaves) , Εξωτερικοί, τερματικοί

Ονομάζονται οι κόμβοι από τους οποίους δεν αρχίζει κάποιο υποδένδρο. Όλοι οι άλλοι λέγονται **κλαδιά** (branches), **εσωτερικοί, μη τερματικοί**.

❖ Η ρίζα ενός δένδρου ονομάζεται **πατέρας** (father) των ριζών των υποδένδρων και οι ρίζες των υποδένδρων ονομάζονται **παιδιά** (children). Οι κόμβοι που έχουν τον ίδιο πατέρα, ονομάζονται **αδέρφια** (brothers).

### ❖ Διατεταγμένο Δένδρο (ordered)

Είναι το δένδρο στο οποίο η διάταξη των κλαδιών έχει σημασία.

### ❖ Βάθος ή Ύψος Δένδρου (η ρίζα βρίσκεται στο επίπεδο 1)

Ονομάζεται το μέγιστο επίπεδο όλων των κόμβων του δένδρου.

## Ορισμοί

### ❖ Δάσος

Είναι ένα σύνολο από περισσότερα του ενός δένδρα ξένα μεταξύ τους.

### ❖ Δενδρόκηπος

Είναι ένα δάσος με διατεταγμένα δένδρα.

### ❖ Τυχαίο Δένδρο

Πρόκειται για το δένδρο που προκύπτει από εισαγωγές κλειδιών με τυχαίες τιμές.

❖ Πλήρες ονομάζεται το δένδρο ( $d$ :βαθμού  $h$ :ύψους) το οποίο έχει το μέγιστο αριθμό κόμβων (όταν όλοι οι κόμβοι του εκτός των κόμβων του τελευταίου επιπέδου έχουν  $d$  υποδένδρα).

## Ορισμοί

### ❖ Μήκος Μονοπατιού Κόμβου (path length)

Ονομάζεται ο αριθμός των κλαδιών που μεσολαβούν μεταξύ του κόμβου αυτού και της ρίζας.

Γενικά ένας κόμβος του  $i$  επιπέδου, έχει μήκος μονοπατιού  $i-1$ .

### ❖ Εσωτερικό Μήκος Μονοπατιού (internal path length - IPL)

Ονομάζεται το άθροισμα των μονοπατιών όλων των κόμβων

$$IPL = \sum_{i=1}^n l_i$$

### ❖ Ζυγισμένο Εσωτερικό Μήκος Μονοπατιού

$$IPL = \sum_{i=1}^n w_i l_i$$

### ❖ Μέσο Εσωτερικό Μήκος Μονοπατιού

$$E[IPL] = \frac{1}{n} \sum_{i=2}^n n_i \times (i-1) \quad n_i: \text{αριθμό κόμβων στο επίπεδο } i.$$

## Ορισμοί

### ❖ Ειδικοί Κόμβοι (special nodes)

Ονομάζονται οι κόμβοι που προσαρτώνται σε ένα δένδρο έτσι ώστε όλοι οι κόμβοι να έχουν τον ίδιο βαθμό, τον βαθμό του δένδρου.

### ❖ Εξωτερικό Μήκος Μονοπατιού (external path length)

Ορίζεται το άθροισμα των μονοπατιών των ειδικών κόμβων.

### ❖ Μέσο Εξωτερικό Μήκος Μονοπατιού

$$E[EPL] = \frac{1}{m} \sum_{i=2}^h m_i \times (i-1) \quad m_i: \text{αριθμός ειδικών κόμβων}$$



## Προτάσεις

### □ Πρόταση 1

Ο αριθμός των κόμβων ενός πλήρους δυαδικού δένδρου ύψους  $h$  είναι  $2^h - 1$ .

### □ Πρόταση 2

Ένα πλήρες δυαδικό δένδρο με  $n$  κόμβους έχει ύψος τουλάχιστον  $\log_2(n+1)$ .

### □ Πρόταση 3

Για κάθε μη κενό δυαδικό δένδρο, αν  $n_0$  είναι ο αριθμός των τερματικών κόμβων και  $n_2$  είναι ο αριθμός των κόμβων βαθμού 2 τότε ισχύει:

$$n_0 = n_2 + 1$$

## Προτάσεις

### □ Πρόταση 4

Ο αριθμός των διαφορετικών δυαδικών δένδρων που μπορούν να παραχθούν με  $n$  κόμβους ισούται με τον αριθμό των τρόπων που μπορούν να πολλαπλασιαστούν  $n+1$  πίνακες, καθώς επίσης ισούται με τον αριθμό των μεταθέσεων που μπορούν να παραχθούν από  $n$  ακεραίους με τη βοήθεια μίας στοίβας.

### □ Πρόταση 5

Ο αριθμός των διακεκριμένων δυαδικών δένδρων με  $n$  κόμβους είναι:

$$b_n \approx \frac{4^n}{n\sqrt{\pi n}}$$

### □ Πρόταση 6

Σε ένα δυαδικό δένδρο με  $n$  κόμβους οι ποσότητες IPL και EPL συνδέονται με τη σχέση

$$\text{EPL} = \text{IPL} + 2n$$

## Δυαδικά Δένδρα

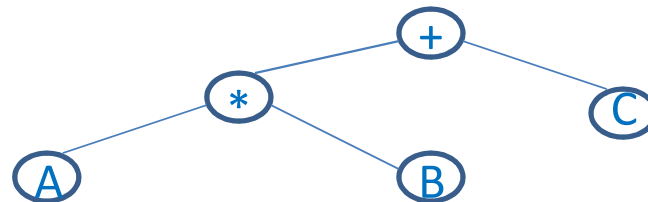
### ❖ Ορισμός

Ένα δυαδικό δένδρο αποτελείται από ένα πεπερασμένο σύνολο κόμβων. Το δένδρο είναι είτε άδειο, είτε αποτελείται από μία ρίζα και δύο άλλα δυαδικά δένδρα που ονομάζονται αριστερό και δεξιό υποδένδρο.

### ❖ Δένδρο Έκφρασης

Είναι το δένδρο το οποίο αναπαριστάει μία αλγεβρική έκφραση.

π.χ.  $A * B + C$



## Δυαδικά Δένδρα / Διάσχιση

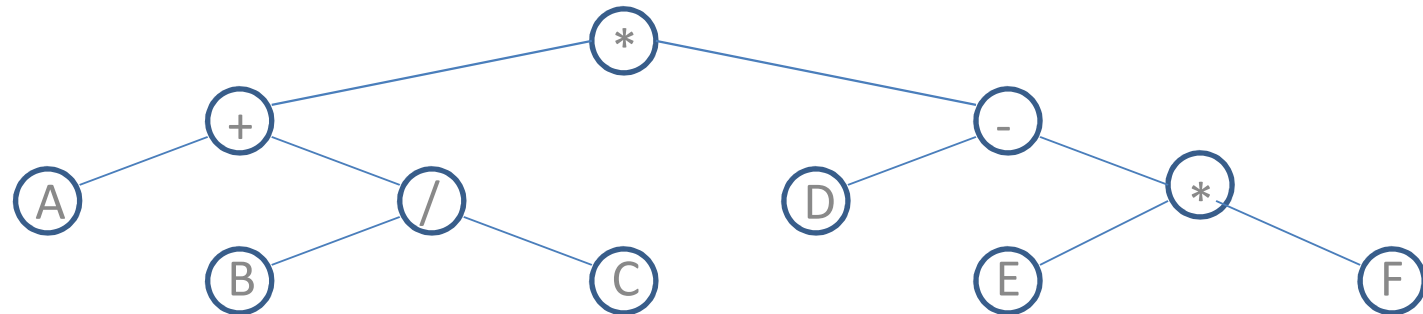
### ❖ Διάσχιση ή Επίσκεψη

Πρόκειται για την λειτουργία της επίσκεψης όλων των κόμβων του δένδρου για μία μόνο φορά τον καθένα.

### ❖ Τρόποι Επίσκεψης

1. **Προδιατεταγμένος (preorder)** – Επισκεπτόμαστε πρώτα τον κόμβο και μετά επισκεπτόμαστε το αριστερό και το δεξιό υποδένδρο. (+\*ABC – προθεματική μορφή - prefix).
2. **Μεταδιατεταγμένος (post-order)** – Επισκεπτόμαστε πρώτα το αριστερό υποδένδρο, μετά το δεξιό και τέλος τον κόμβο. (AB\*C+ - επιθεματική μορφή).
3. **Ενδοδιατεταγμένη (in-order)** – Επισκεπτόμαστε πρώτα το αριστερό υποδένδρο, μετά τον κόμβο και τέλος το δεξιό υποδένδρο. (A\*B+C).

## Διαδικά Δένδρα / Αποθήκευση



### 1ος Τρόπος (Καταγράφουμε τα παιδιά των κόμβων)

1	2	3	4	5	6	7	8	9	10	11
*	+	-	/	*	A	B	C	D	E	F
2	6	9	7	10	0	0	0	0	0	0
3	4	5	8	11	0	0	0	0	0	0

### 2ος Τρόπος (Καταγράφουμε τον πατέρα των κόμβων)

1	2	3	4	5	6	7	8	9	10	11
*	+	-	/	*	A	B	C	D	E	F
0	1	1	2	3	2	4	4	3	5	5

## Δυαδικά Δένδρα / Αποθήκευση

### ❖ Διάσχιση ή Επίσκεψη

Πρόκειται για την λειτουργία της επίσκεψης όλων των κόμβων του δένδρου για μία μόνο φορά τον καθένα.

3<sup>ος</sup> Τρόπος (Πλήρη δυαδικά δένδρα) Αποθήκευση από επάνω προς τα κάτω και από αριστερά προς τα δεξιά.

- 1) Ο πατέρας του κόμβου  $i$  βρίσκεται στη θέση  $\lfloor i/2 \rfloor$  αν  $i=1$  τότε το  $i$  είναι η ρίζα και δεν έχει πατέρα.
- 2) Το αριστερό παιδί βρίσκεται στη θέση  $2i$ , αν  $2i \leq n$ . Αν  $2i > n$  τότε ο κόμβος  $i$  δεν έχει αριστερό παιδί.
- 3) Το δεξιό παιδί βρίσκεται στη θέση  $2i+1$ , αν  $2i+1 \leq n$ . Αν  $2i+1 > n$  τότε ο κόμβος  $i$  δεν έχει δεξιό παιδί.

## Δυαδικά Δένδρα Αναζήτησης

❖ Πρόκειται για δυαδικά δένδρα τα οποία έχουν ένα κλειδί συσχετισμένο με καθέναν από τους εσωτερικούς κόμβους, με την πρόσθετη ιδιότητα ότι το κλειδί οποιουδήποτε κόμβου είναι μεγαλύτερο (ή ίσο) από τα κλειδιά όλων των κόμβων του αριστερού υποδένδρου αυτού του κόμβου και μικρότερο (ή ίσο) από τα κλειδιά των κόμβων του δεξιού υποδένδρου αυτού του κόμβου.

❖ Η αναζήτηση γίνεται ως εξής:

1. Αν το δένδρο είναι άδειο, έχουμε μία ανεπιτυχή αναζήτηση.
2. Αν το κλειδί αναζήτησης είναι ίσο με το κλειδί της ρίζας τότε έχουμε μια επιτυχή αναζήτηση.
3. Αλλιώς, ψάχνουμε (αναδρομικώς) στο κατάλληλο κάθε φορά υποδένδρο, ανάλογα με την τιμή του κλειδιού αναζήτησης.

## Δυαδικά Δένδρα Αναζήτησης

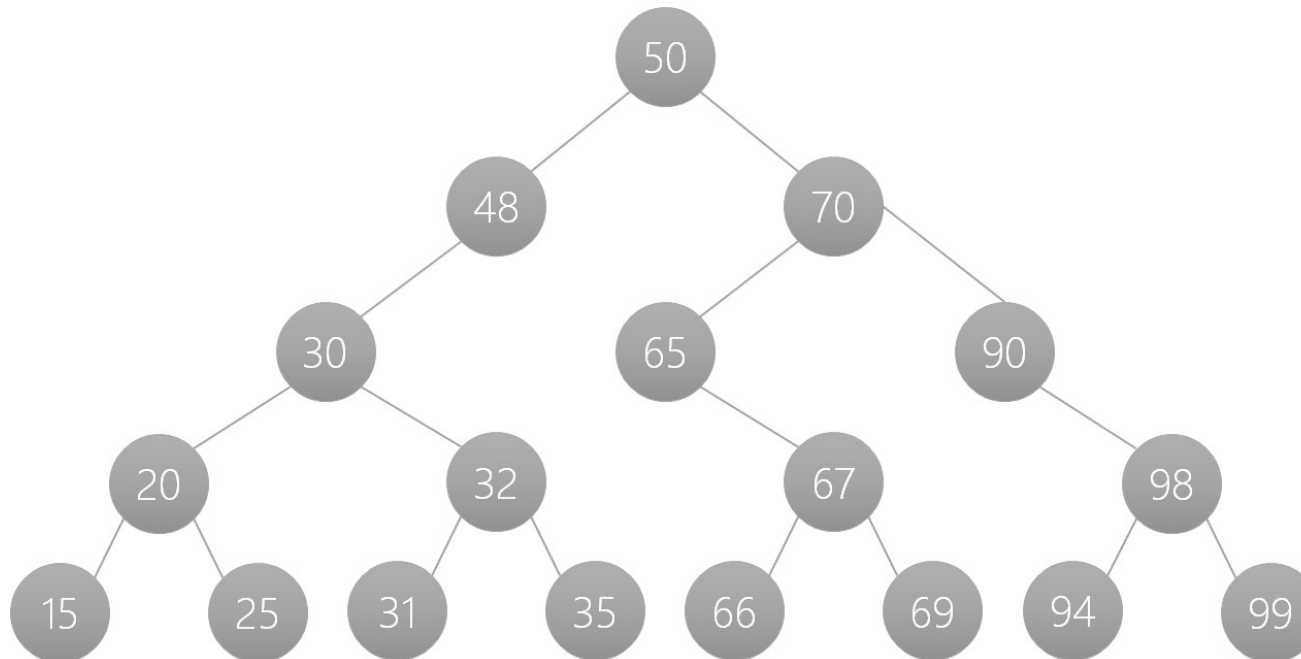
❖ Μεγάλο ενδιαφέρον παρουσιάζουν και οι διαδικασίες της **διαγραφής** και **εισαγωγής** κόμβων στο δένδρο.

- 1) Ενώ η διαδικασία της εισαγωγής είναι απλή, η διαδικασία της διαγραφής είναι απλή μόνον όταν ο κόμβος έχει έναν μόνο απόγονο.
- 2) Στην αντίθετη περίπτωση το διαγραφόμενο κλειδί πρέπει να αντικατασταθεί είτε από το κλειδί του πιο δεξιού κόμβου του αριστερού υποδένδρου, είτε από κλειδί του πιο αριστερού κόμβου του δεξιού υποδένδρου.



# Δυαδικά Δένδρα Αναζήτησης

## Παράδειγμα



## Δυαδικά Δένδρα Αναζήτησης

### □ Πρόταση 7

Η μέση τιμή του αριθμού των συγκρίσεων για επιτυχή αναζήτηση σε ένα πλήρες δυαδικό δένδρο είναι:

$$E \approx \log \frac{n}{2}$$

### □ Πρόταση 8

Η μέση τιμή του αριθμού των συγκρίσεων για την επιτυχή και την ανεπιτυχή αναζήτηση σε ένα τυχαίο δυαδικό δένδρο συνδέονται με την ταυτότητα:

$$E = A \times \frac{n+1}{n} - 1$$

## Δυαδικά Δένδρα Αναζήτησης

### □ Πρόταση 9

Η μέση τιμή του αριθμού των συγκρίσεων για την ανεπιτυχή αναζήτηση σε τυχαίο δυαδικό δένδρο είναι :

$$A = 2 \times H_{n+1} - 2$$

$$H_n = g + \ln n + \frac{1}{2n} - \frac{1}{12n^2}$$

$$g = 0.577 \text{ Euler constant}$$

### □ Πρόταση 10

Η μέση τιμή του αριθμού των συγκρίσεων για την επιτυχή και την ανεπιτυχή αναζήτηση σε τυχαίο δυαδικό δένδρο είναι:

$$A = 1.386 \times \log n - 0.846$$

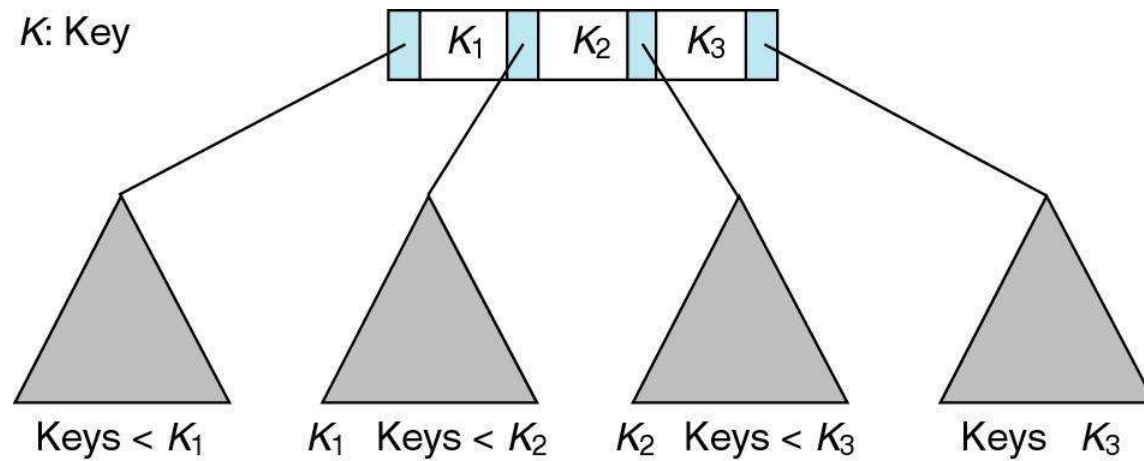
$$E = 1.386 \times \log n - 1.846$$

## Πολυκατευθυνόμενα Δένδρα (Δένδρα $m$ -δρόμων)

### ❖ Ορισμός - Ιδιότητες

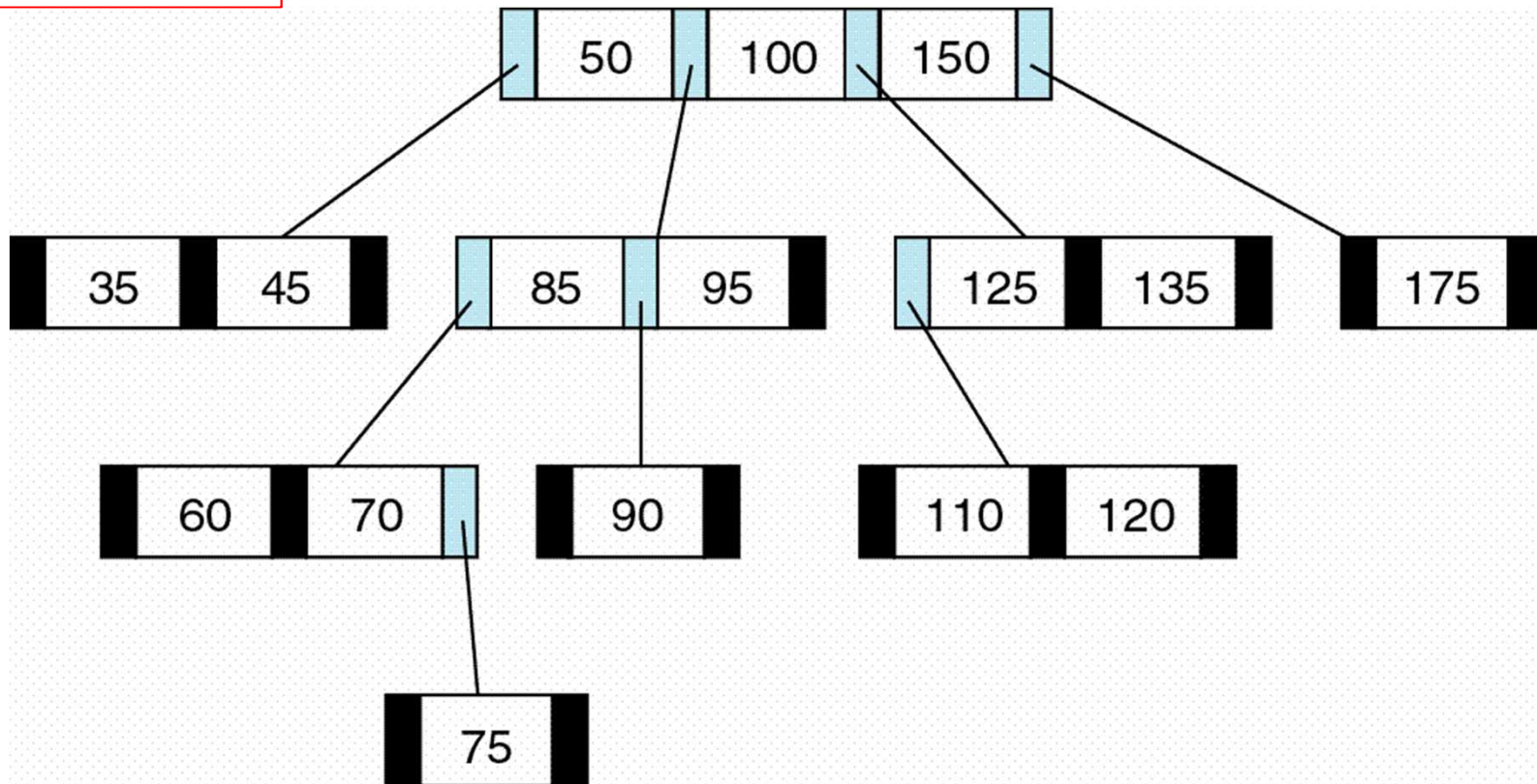
- 1) Στο αντίστοιχο εκτεταμένο δένδρο αναζήτησης (που λαμβάνεται αντικαθιστώντας τους μηδενικούς δείκτες με εξωτερικούς κόμβους), κάθε εσωτερικός κόμβος έχει μέχρι και  $m$  παιδιά και από  $1$  μέχρι  $m-1$  στοιχεία.
- 2) Κάθε κόμβος με  $p$  στοιχεία έχει ακριβώς  $p+1$  παιδιά.
- 3) Ας θεωρήσουμε έναν κόμβο με  $p$  στοιχεία. Έστω  $k_1, k_2, k_3 \dots k_p$  τα κλειδιά αυτών των στοιχείων. Τα στοιχεία τοποθετούνται στη σειρά ώστε  $k_1 < k_2 < k_3 \dots < k_p$ . Έστω  $c_0, c_1, c_2, \dots c_p$  τα  $p+1$  παιδιά του κόμβου. Τα στοιχεία στο υποδένδρο με ρίζα  $c_0$  έχουν κλειδιά μικρότερα από  $k_1$ , αυτά στο υποδένδρο με ρίζα  $c_p$  έχουν κλειδιά μεγαλύτερα από  $k_p$  και αυτά στο υποδένδρο με ρίζα  $c_i$  έχουν κλειδιά μεγαλύτερα από  $k_i$ , αλλά μικρότερα από  $k_{i+1}$ , για  $1 \leq i < p$ .

## Πολυκατευθυνόμενα Δένδρα (Δένδρα $m$ -δρόμων)



## Πολυκατευθυνόμενα Δένδρα (Δένδρα $m$ -δρόμων)

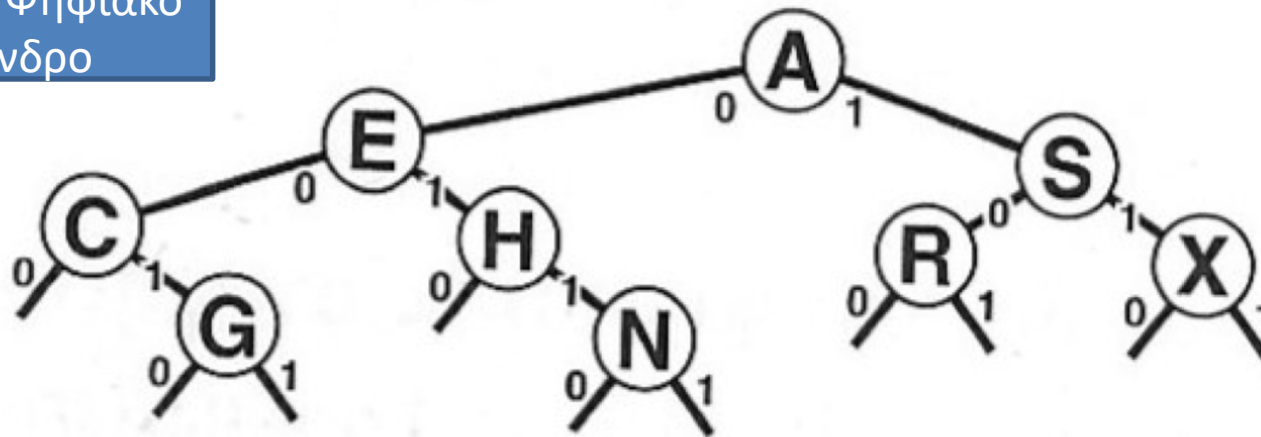
Παράδειγμα  $m=4$



# Ψηφιακά Δένδρα Αναζήτησης

Πρόκειται για δέντρα στα οποία η διάσχιση γίνεται όχι με συγκρίσεις των κλειδιών αλλά με συγκρίσεις επιλεγμένων bit των κλειδιών.

Τυπικό Ψηφιακό  
Δένδρο



A:00001  
E:00101  
S:10011  
C:00011  
H:01000  
R:10010  
X:11000  
G:00111  
N:01110

Ερώτηση: Το κλειδί **M:01101** που θα εισαχθεί ?

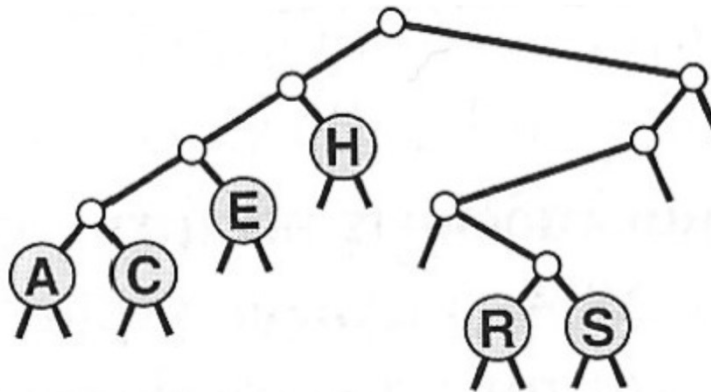
# Ψηφιακά Δένδρα Αναζήτησης

Πρόκειται για ψηφιακά δένδρα στα οποία τα κλειδιά βρίσκονται μόνον σε φύλλα του δένδρου.

Trie Ψηφιακό  
Δένδρο

**Μειονεκτήματα:**

- (1) Άχρηστοι κόμβοι
- (2) Δύο τύποι κόμβων



A:00001  
E:00101  
S:10011  
C:00011  
H:01000  
R:10010

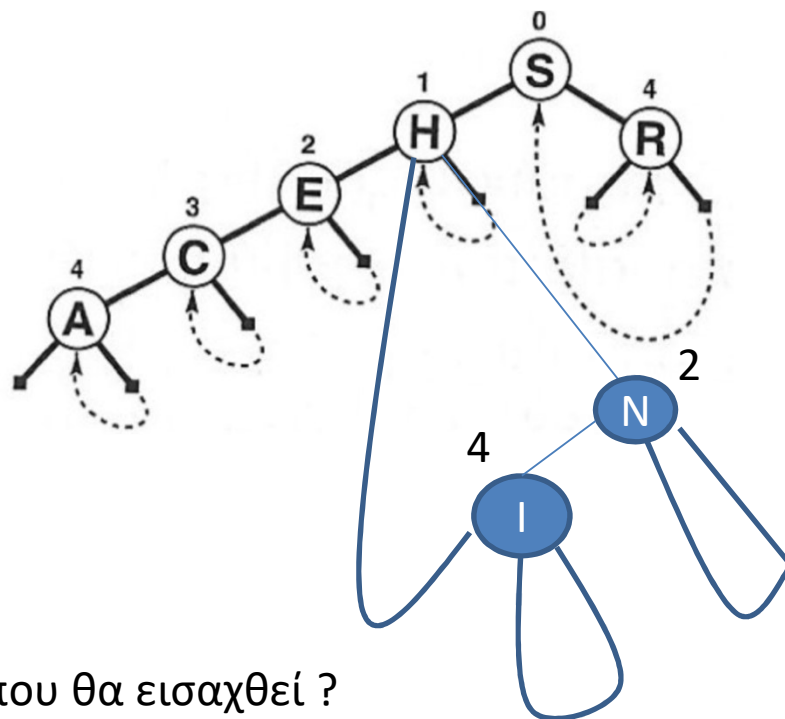
**Ερώτηση:** Το **I:01001** που θα εισαχθεί ?



## Μη Δυαδικά Δένδρα Αναζήτησης / Δένδρα Patricia (Practical Algorithm to Retrieve Information Coded in Alphanumeric)

Πρόκειται για ψηφιακά δένδρα στα οποία η διάσχιση γίνεται σύμφωνα με ένα προκαθορισμένο αριθμοδείκτη του bit κάθε φορά.

Patricia Trie  
Ψηφιακό Δένδρο



A:00001  
E:00101  
S:10011  
C:00011  
H:01000  
R:10010

N=01110

Ερώτηση: Το **I:01001** που θα εισαχθεί ?