

Αρχιτεκτονικές Πληροφοριακών Συστημάτων

Δομές δεδομένων
Οργάνωση αρχείων

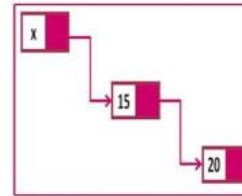
Δομές Δεδομένων

Δομή Δεδομένων (data structure)

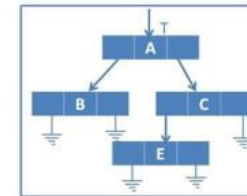
- Δομή δεδομένων είναι μια συλλογή δεδομένων που έχουν μεταξύ τους μια συγκεκριμένη σχέση
- Παραδείγματα δομών δεδομένων
 - Πίνακες (διανύσματα - συστοιχίες)
 - Εγγραφές
 - Συνδεδεμένες λίστες
 - Δένδρα
 - Γράφοι



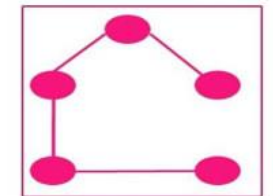
Sorting



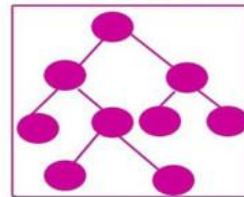
Link list



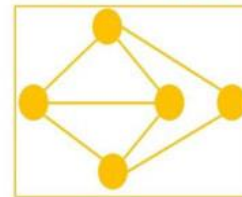
list



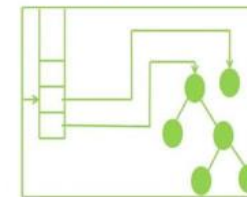
spanning tree



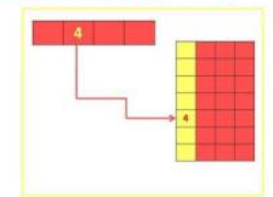
Tree



Graph



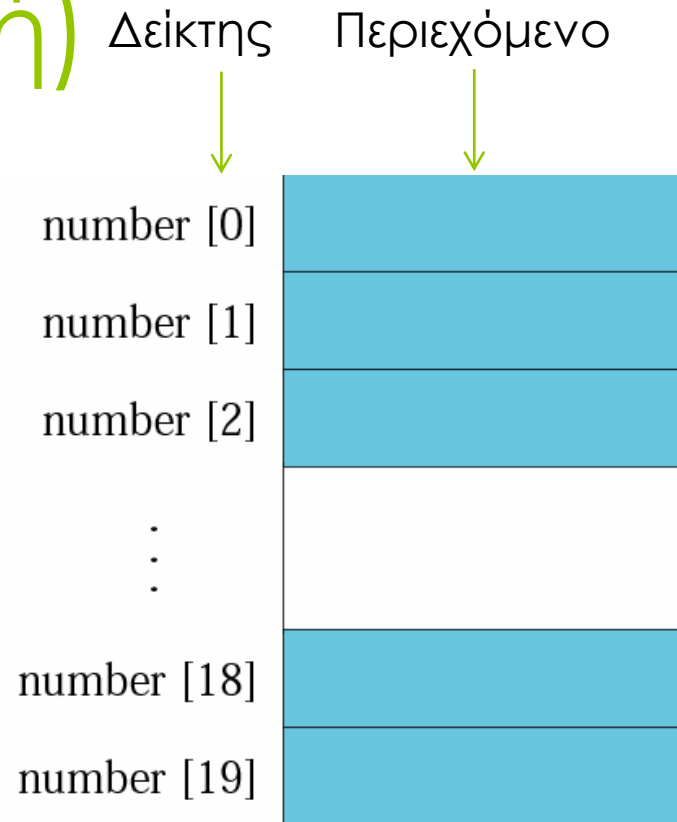
Stack



Hashing

Πίνακες (*στατική δομή)

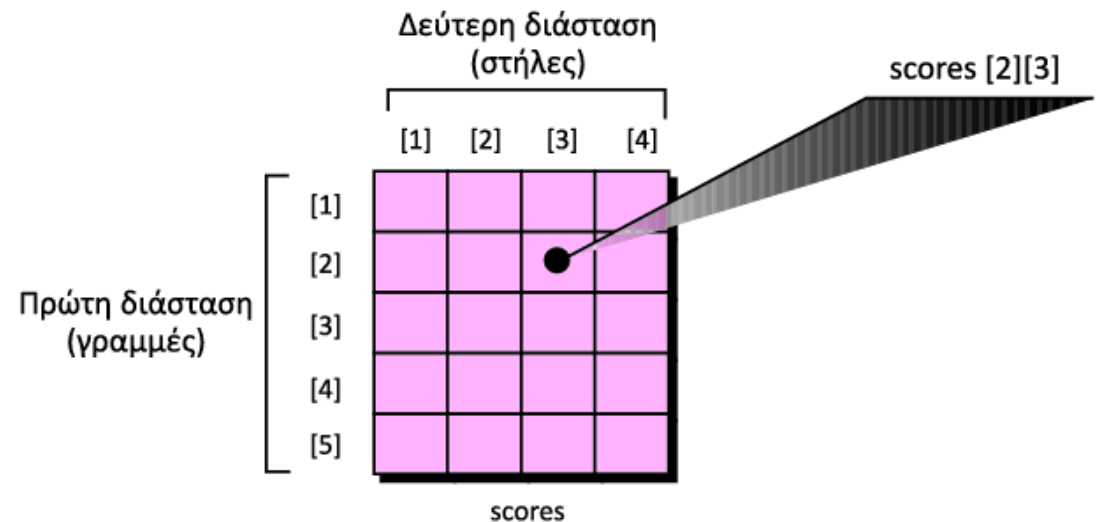
- Πίνακας είναι **μια σειριακή δομή στοιχείων του ίδιου τύπου** στην οποία η αναφορά σε κάθε στοιχείο γίνεται με ένα δείκτη που υποδηλώνει την θέση του στοιχείου
- Στην γλώσσα C και σε πολλές άλλες γλώσσες το πρώτο στοιχείο ενός μονοδιάστατου πίνακα έχει ως δείκτη την τιμή 0



Πολυδιάστατοι πίνακες

- Υπάρχουν **πίνακες πολλών διαστάσεων** αλλά συνήθως χρησιμοποιούνται μόνο οι **δισδιάστατοι**
- Οι δισδιάστατοι πίνακες έχουν **γραμμές και στήλες**
- Ένας πίνακας 5X4 έχει 5 γραμμές και 4 στήλες
- Συνήθως χρησιμοποιούνται ένθετες εντολές επανάληψης for για να αναφερθούμε στα στοιχεία του πίνακα

```
for (int i=0;i<5;i++)  
    for (int j=0;j<4;j++)  
        {...a[i][j]...}
```

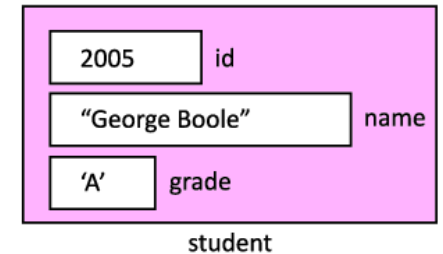
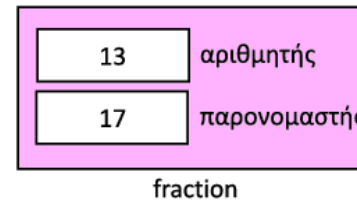


Βασικές λειτουργίες σε πίνακες

- Αναζήτηση
 - Ταξινόμηση
 - Άθροισμα στοιχείων
 - Εύρεση Μεγίστου
 - Συγχώνευση
 - Διαχωρισμός
- Παρόλο που η αναζήτηση, η ανάκτηση, και η διάσχιση των στοιχείων ενός πίνακα γίνονται εύκολα, **η εισαγωγή και η διαγραφή είναι χρονοβόρες.**
 - Αυτό συμβαίνει επειδή τα στοιχεία πρέπει να μετατοπίζονται προς τα κάτω πριν από μια εισαγωγή και προς τα πάνω μετά από μια διαγραφή

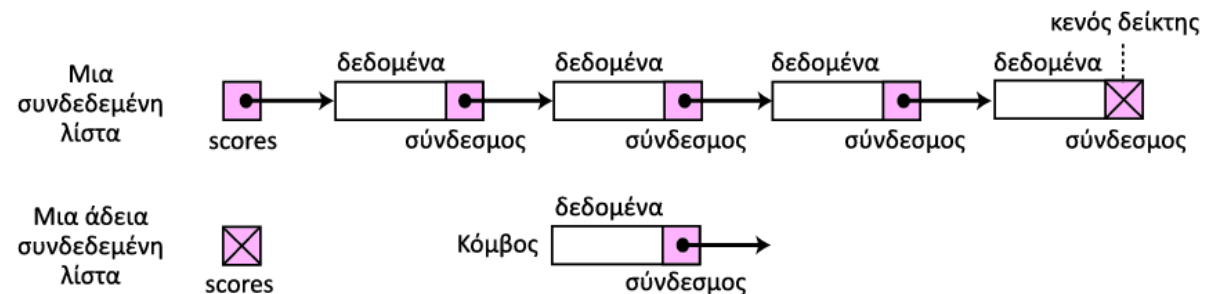
Εγγραφές (records)

- Μια **εγγραφή** είναι μια **συλλογή από σχετικά μεταξύ τους στοιχεία, πιθανώς διαφορετικών τύπων**, η οποία έχει ένα μοναδικό όνομα
- Κάθε στοιχείο μιας εγγραφής ονομάζεται πεδίο
- Μια μεταβλητή μπορεί να λάβει ως τύπο δεδομένων τον τύπο της εγγραφής
- Στο πρώτο παράδειγμα η εγγραφή fraction έχει δύο πεδία, τα οποία είναι και τα δύο ακέραιοι.
- Στο δεύτερο παράδειγμα η εγγραφή student έχει τρία πεδία τα οποία ανήκουν σε δύο διαφορετικούς τύπους.



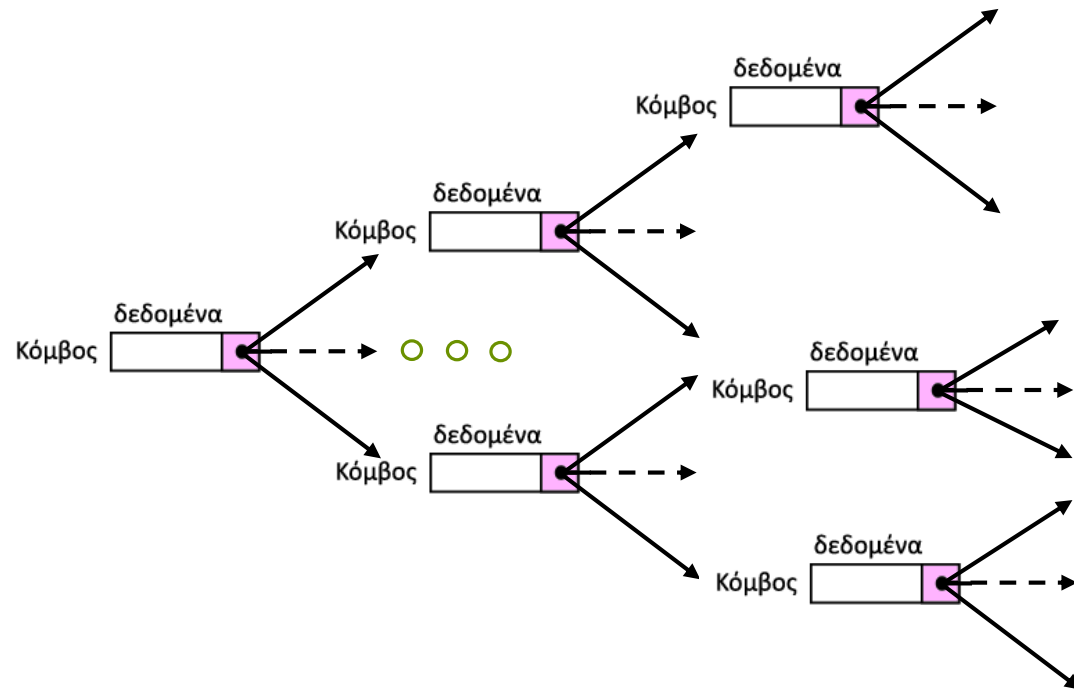
Συνδεδεμένες λίστες (*δυναμική δομή)

- Μια **συνδεδεμένη λίστα** είναι μια συλλογή δεδομένων στην οποία **κάθε στοιχείο «γνωρίζει» την θέση του επόμενου** του
- Χρησιμοποιούνται για την **αποθήκευση δεδομένων** στα οποία πρόκειται να γίνουν πολλές **εισαγωγές** και **διαγραφές**
- Δεν αποτελεί την καλύτερη λύση όταν πρέπει να εκτελούνται συχνά αναζητήσεις



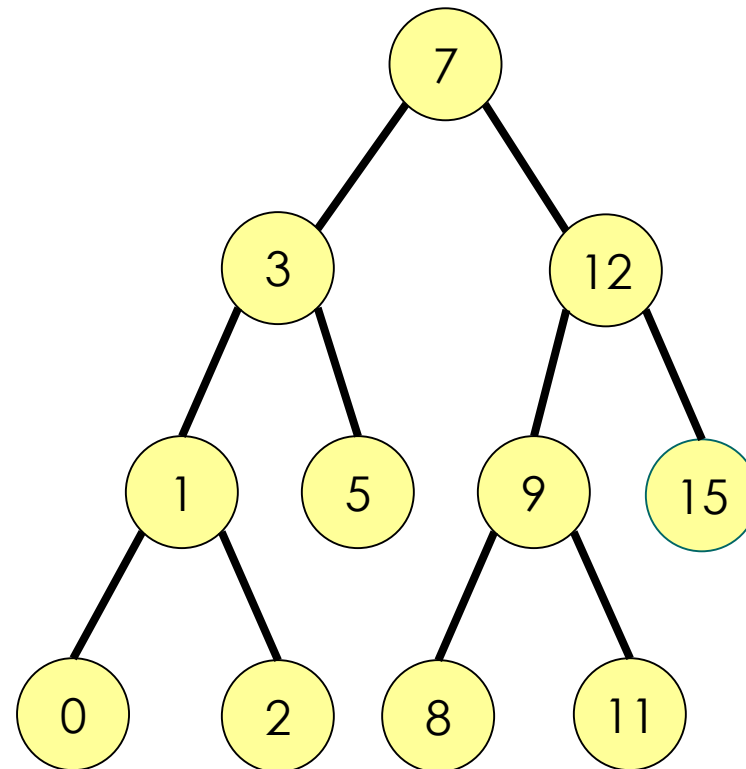
Δένδρα

- Γενίκευση της διασυνδεδεμένης λίστας
- Κάθε κόμβος ξέρει (μπορεί να «δείχνει» τη θέση Κ άλλων κόμβων)



Ταξινομημένα Δυαδικά Δένδρα

- Ένας διαφορετικός τρόπος αναπαράστασης μιας ταξινομημένης λίστας αντικειμένων.
- Κάθε κόμβος έχει το πολύ 2 «παιδιά».
- Κάθε κόμβος στα αριστερά έχει μικρότερη τιμή από κάθε κόμβου στα δεξιά του.
- Υποστηρίζει τις παρακάτω διεργασίες σε γρήγορο χρόνο ($O(\log n)$) πολυπλοκότητα μέσης περίπτωσης:
 - Αναζήτηση ενός υπάρχοντος αντικειμένου.
 - Εισαγωγή ενός καινούργιου αντικειμένου, αν δεν είναι ήδη παρόν
- Υποστηρίζει την εκτύπωση όλων των αντικειμένων σε $O(n)$ χρόνο
- Η εισαγωγή ενός αντικειμένου σε μια συνεχόμενη ακολουθία είναι πιο ακριβή ($O(n)$ στη χειρότερη περίπτωση)

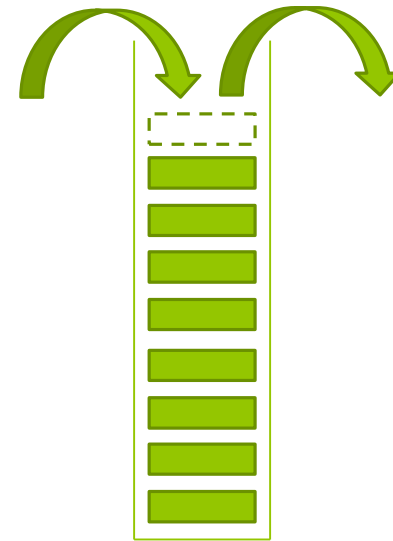


Στοιβες και ουρές

- Ουρά
Δομή FIFO (First In First Out)



- Στοιβα
Δομή LIFO (Last In First Out)



Σε ουρές και στοιβες σταθερού μήκους, ο ρυθμός των εισερχομένων και των εξερχομένων είναι ίδιος

Αρχεία και φάκελοι

Σύστημα αρχείων

- Στην βοηθητική μνήμη τα δεδομένα αποθηκεύονται με την **μορφή αρχείων**
- Αρχείο **είναι μια συλλογή δεδομένων που έχουν κάποια σχέση μεταξύ τους** και στα οποία αποδίδεται ένα **όνομα** και μια **θέση** αποθήκευσης
- Η **κατάληξη** του αρχείου δείχνει τον τύπο του (txt, mp3, jpg, doc, avi, ppt, xls)
- **Κατάλογοι** – υποκατάλογοι βοηθούν στην αποδοτικότερη οργάνωση των αρχείων

Αποθήκευση αρχείων

- Βοηθητικές αποθηκευτικές συσκευές (Μόνιμη αποθήκευση)
 - Μαγνητική Ταινία
 - Μαγνητικός Δίσκος
 - Οπτικός Δίσκος
 - Μνήμες τύπου FLASH (USB-sticks)
 - Δίσκοι SSD

Τύποι αρχείων με βάση το περιεχόμενο

Αρχεία κειμένου

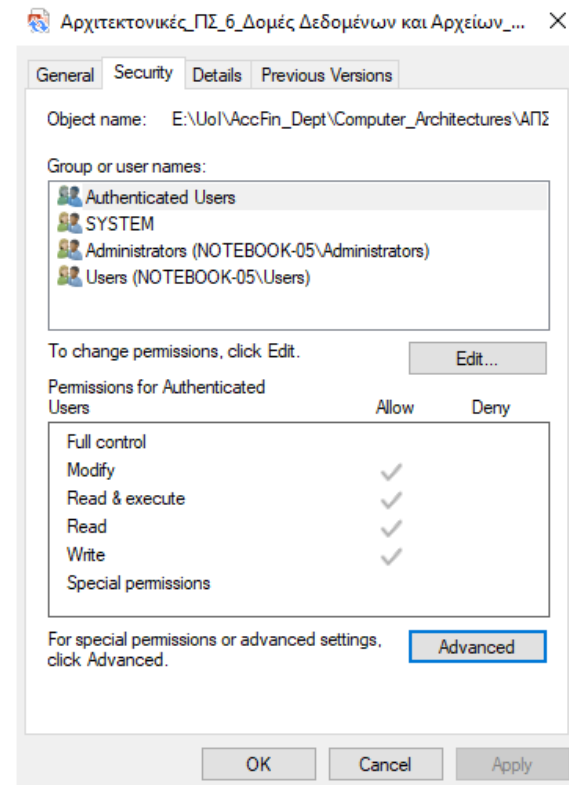
- Ένα αρχείο κειμένου είναι ένα αρχείο χαρακτήρων
- Αν περιέχουν αριθμούς αυτοί αποθηκεύονται με κωδικοποίηση χαρακτήρων (ASCII, Unicode) και όχι στην εσωτερική μορφή αναπαράστασης αριθμητικών δεδομένων (συμπλήρωμα ως προς 2)

Δυαδικά αρχεία

- Ένα δυαδικό αρχείο είναι μια συλλογή δεδομένων αποθηκευμένων στην εσωτερική μορφή του υπολογιστή
- Τα δεδομένα ενός δυαδικού αρχείου κειμένου έχουν νόημα μόνο όταν ερμηνεύονται κατάλληλα από ένα πρόγραμμα

Λειτουργίες σε αρχεία

- **Δημιουργία** αρχείου
- **Διαγραφή** αρχείου
- **Μετονομασία** αρχείου
- **Αντιγραφή** αρχείου
- **Άνοιγμα** αρχείου
- **Κλείσιμο** αρχείου
- **Διάβασμα** δεδομένων αρχείου
- **Ενημέρωση** δεδομένων αρχείου
- **Διαγραφή** δεδομένων αρχείου
- **Εισαγωγή** νέων δεδομένων σε αρχείο

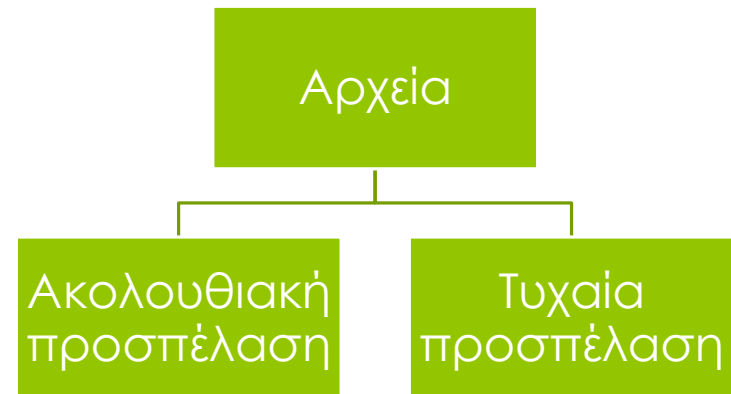


Αρχεία δεδομένων

- Τα αρχεία δεδομένων αποτελούνται από εγγραφές (π.χ. αρχείο πελατών με κάθε εγγραφή τα στοιχεία ενός πελάτη)
- Κάθε εγγραφή αποτελείται από πεδία (κάθε εγγραφή πελάτη αποτελείται από πεδία όπως κωδικός, όνομα, διεύθυνση κ.α.)
- Πρωτεύον κλειδί είναι ένα πεδίο (ή ο συνδυασμός περισσοτέρων πεδίων) που χαρακτηρίζει μοναδικά μια εγγραφή

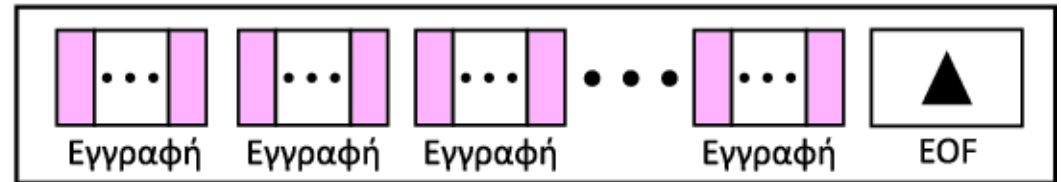
Μέθοδοι προσπέλασης αρχείων δεδομένων

- Η **μέθοδος προσπέλασης** καθορίζει τον τρόπο με τον οποίο προσπελάζονται τα δεδομένα που περιέχονται σε ένα αρχείο



Ακολουθιακά αρχεία

- Οι εγγραφές ενός αρχείου προσπελάζονται σειριακά δηλαδή η μια μετά την άλλη από την αρχή προς το τέλος
- Μετά την τελευταία εγγραφή υπάρχει ένα σημάδι τέλους αρχείου (EOF=End Of File)
- Δεν υπάρχει διεύθυνση για κάθε εγγραφή
- Είναι κατάλληλα για εφαρμογές στις οποίες απαιτείται η προσπέλαση σε όλες τις εγγραφές του αρχείου
- Δεν είναι αποδοτικά όταν απαιτείται τυχαία προσπέλαση
- Η αποθήκευση αρχείων σε **μαγνητική ταινία** είναι το χαρακτηριστικότερο παράδειγμα στο οποίο έχουμε **σειριακή προσπέλαση δεδομένων**

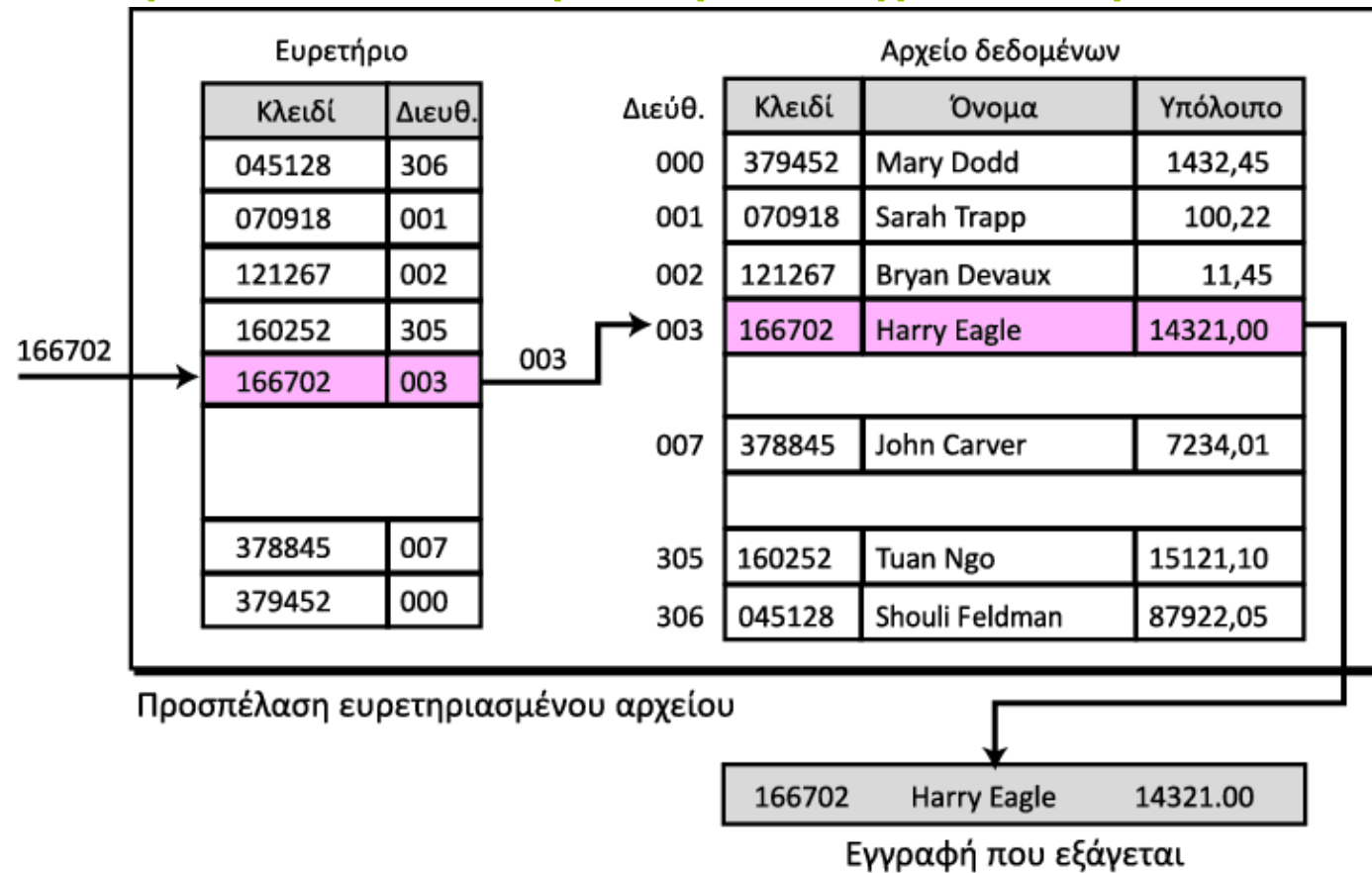


Ακολουθιακό αρχείο

Ευρετηριασμένα αρχεία (indexed files)

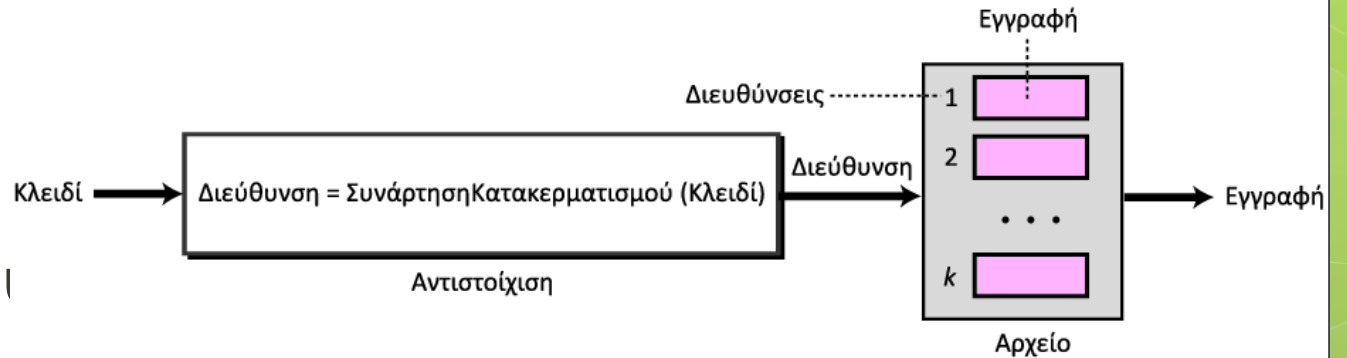
- Ένα ευρετηριασμένο αρχείο αποτελείται από ένα **αρχείο δεδομένων** και ένα **ευρετήριο**
- Το **ευρετήριο** είναι ένα μικρό αρχείο με δύο πεδία: το **κλειδί** της εγγραφής και την αντίστοιχη **διεύθυνση** της
- Το **αρχείο δεδομένων** είναι ένα **ακολουθιακό αρχείο** που περιέχει τις εγγραφές
- Προσπέλαση
 - **Φορτώνεται** στην μνήμη όλο το αρχείο **ευρετήριο**
 - **Ερευνώνται οι εγγραφές του ευρετηρίου** με έναν αποδοτικό αλγόριθμο αναζήτησης
 - **Ανακτάται η διεύθυνση της εγγραφής**
 - Χρησιμοποιείται η **διεύθυνση για να ανακτηθεί η κατάλληλη εγγραφή δεδομένων**

Αναπαράσταση ευρετηριασμένου αρχείου



Κατακερματισμένα αρχεία (Hashed files)

- Στα **κατακερματισμένα** αρχεία δεν υπάρχει ευρετήριο
- Τον ρόλο του ευρετηρίου τον αναλαμβάνει **μια συνάρτηση** που δέχεται ένα κλειδί και το **αντιστοιχεί σε μια διεύθυνση**



Άμεσος κατακερματισμός

- Το **κλειδί** αποτελεί και την **διεύθυνση** του αρχείου δεδομένων
- **Δεν είναι αποδοτικός τρόπος διαχείρισης χώρου**
 - Π.χ. αν το ΑΦΜ (9 ψηφία) είναι το κλειδί τότε θα υπάρχουν εγγραφές με κωδικούς από 000000000 έως 999999999 άρα θα χρειαστεί ένα τεράστιο αρχείο και ένα πολύ μικρό ποσοστό του θα περιέχει πραγματικές εγγραφές

Κατάλογοι

- Κατάλογοι (φάκελοι): Δίνουν την **δυνατότητα οργάνωσης των αρχείων** σε δομή **n-αδικού δένδρου**
- Αναπαρίσταται ως ένας ειδικός τύπος αρχείου
- Μπορεί να περιέχει και άλλες πληροφορίες πέρα από τα αρχεία που περιέχει (π.χ. δικαιώματα πρόσβασης στα περιεχόμενά του)

Κατάλογοι στα ΛΣ.

Το παράδειγμα του UNIX

- Κατάλογος root (/). Βρίσκεται στο υψηλότερο επίπεδο ιεραρχίας του συστήματος αρχείων
- Κατάλογος home. Είναι ο κατάλογος στον οποίο μεταφέρεται ο χρήστης μόλις συνδέεται στο σύστημα. Κάθε χρήστης έχει τον δικό του home κατάλογο
- Τρέχον κατάλογος. Είναι ο κατάλογος στον οποίο βρίσκεται ο χρήστης την τρέχουσα χρονική στιγμή

- Διαδρομή
 - Απόλυτη διαδρομή στο αρχείο file3 **/usr/staff/eleni/αρχείο3**
 - Σχετική διαδρομή στο αρχείο file3 αν ο τρέχων κατάλογος είναι ο **staff** είναι η **eleni/αρχείο3**

