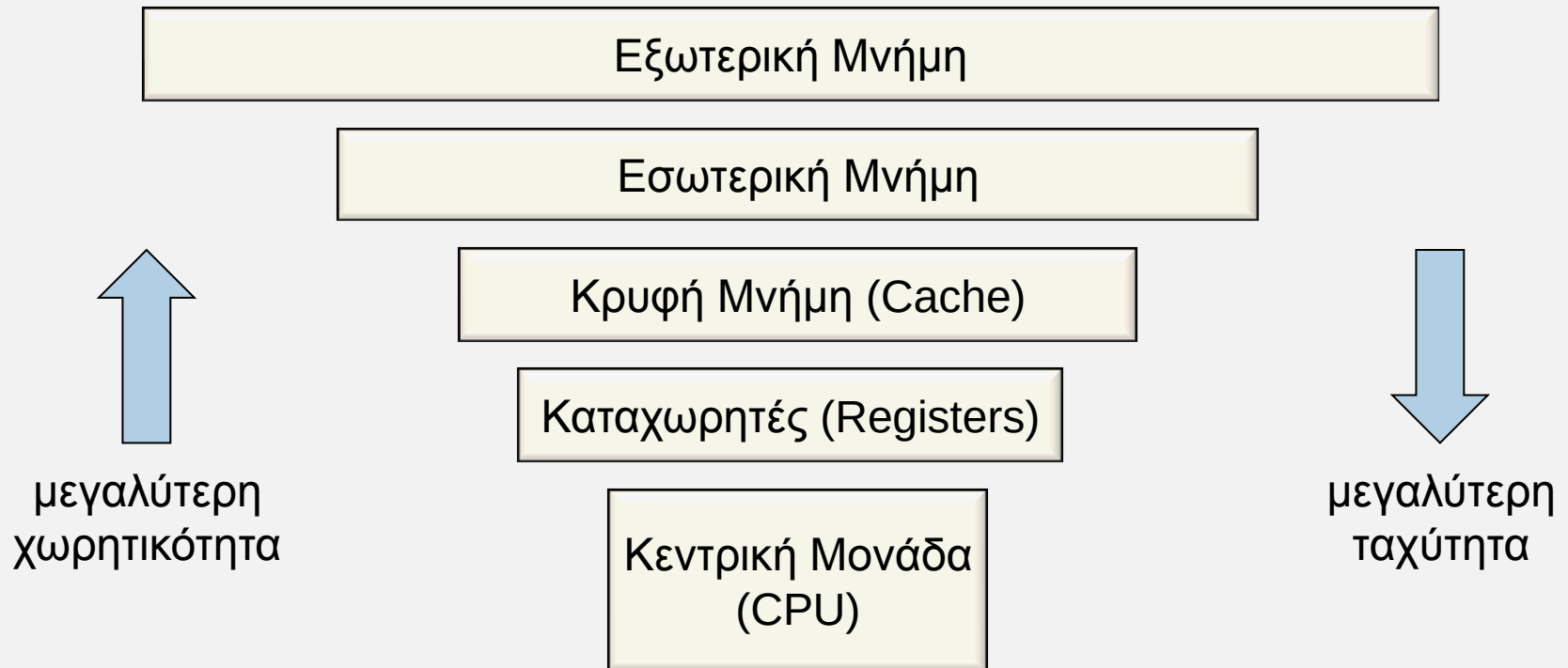


Εξωτερική Αναζήτηση

Ιεραρχία Μνήμης Υπολογιστή



Εξωτερική Αναζήτηση

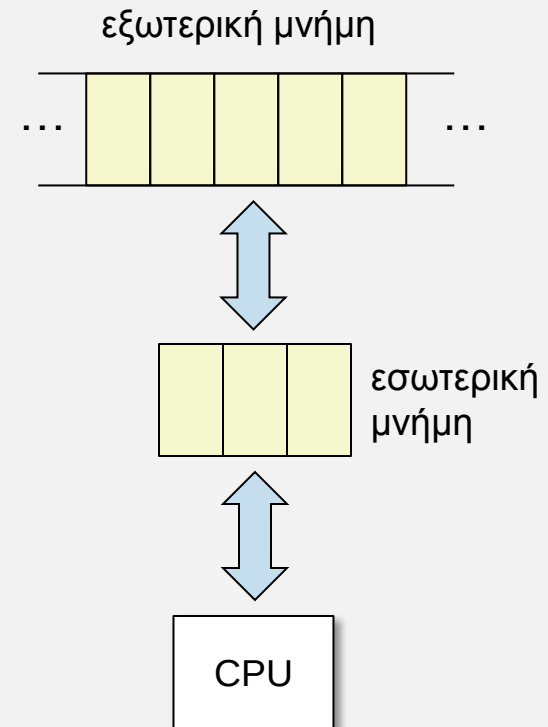
Πολλές σημαντικές εφαρμογές διαχειρίζονται ένα μεγάλο όγκο δεδομένων που είναι αποθηκευμένα σε εξωτερική μνήμη.

Στις περιπτώσεις αυτές πρέπει να λάβουμε υπόψη το χρόνο που απαιτείται για τη μεταφορά δεδομένων από/προς την εξωτερική μνήμη.

Μοντέλο Εξωτερικής Μνήμης

Απλοποιημένο μοντέλο ανάλυσης αλγορίθμων :

Η εξωτερική μνήμη διαιρείται σε σελίδες και μία σελίδα περιέχει ένα μεγάλο αριθμό δεδομένων. Θεωρούμε ότι ο χρόνος των λειτουργιών εισόδου/εξόδου που απαιτείται για την ανάγνωση μιας σελίδας από την εξωτερική μνήμη είναι πολύ μεγαλύτερος από το χρόνο επεξεργασίας των δεδομένων της σελίδας.



Εξωτερική Αναζήτηση

Μοντέλο Εξωτερικής Μνήμης

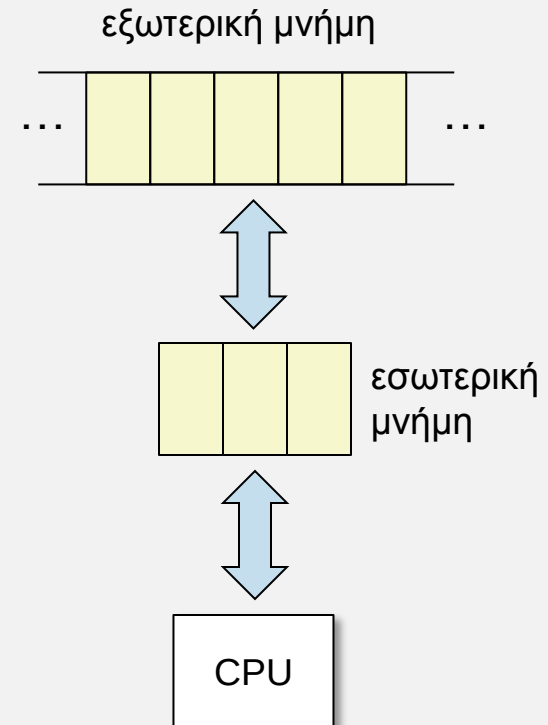
Απλοποιημένο μοντέλο ανάλυσης αλγορίθμων :

Η εξωτερική μνήμη διαιρείται σε σελίδες και μία σελίδα περιέχει ένα μεγάλο αριθμό δεδομένων. Θεωρούμε ότι ο χρόνος των λειτουργιών εισόδου/εξόδου που απαιτείται για την ανάγνωση μιας σελίδας από την εξωτερική μνήμη είναι πολύ μεγαλύτερος από το χρόνο επεξεργασίας των δεδομένων της σελίδας.

Διερεύνηση : Η πρώτη πρόσβαση σε μία σελίδα

Εξωτερική Αναζήτηση : Μας ενδιαφέρουν δομές αναζήτησης που απαιτούν λίγες διερευνήσεις

Μια τυχαία προσπέλαση στην εξωτερική μνήμη είναι πολύ ακριβή σε σχέση με μια τυχαία προσπέλαση στην εσωτερική μνήμη (π.χ. 100000 φορές πιο αργή σε τυπικά συστήματα).



Εξωτερική Αναζήτηση

Ακολουθιακή δομή αρχείου με ευρετήριο

Τα κλειδιά βρίσκονται σε ακολουθιακή διάταξη στις σελίδες. Το ευρετήριο οδηγεί προς το μικρότερο κλειδί κάθε σελίδας.

Για να προσθέσουμε ένα νέο κλειδί πρέπει να ανακατασκευάσουμε ολόκληρη τη δομή.

κλειδιά

706
176
601
153
513
773
742
373
524
766
275
737
574
434
641
207
001
277
061
736
526
562
017
107
147

ευρετήριο

001	
153	
373	
562	
737	

001	
017	
061	
107	
147	

153	
176	
207	
275	
277	

373	
434	
513	
524	
526	

562	
574	
601	
706	
736	

737	
741	
742	
766	
773	

Εξωτερική Αναζήτηση

B-δένδρα

Ένα (a,b) -δένδρο είναι δένδρο αναζήτησης πολλαπλής διακλάδωσης όπου:

- $a \geq 2$ και $b > a$ (συνήθως $b \geq 2 \cdot a$)
- η ρίζα έχει $d-1$ κλειδιά και d παιδιά, $2 \leq d \leq b$
- οι υπόλοιποι εσωτερικοί κόμβοι έχουν $t-1$ κλειδιά και t παιδιά, $a \leq t \leq b$
- οι κενοί κόμβοι (φύλλα) ισαπέχουν από τη ρίζα (δηλ. βρίσκονται στο ίδιο επίπεδο)

B-δένδρο (βαθμού m) : Αντιστοιχεί σε (a,b) -δένδρο με $a = \lceil m/2 \rceil, b = m$

Το ύψος του δένδρου με n κλειδιά είναι μεταξύ $\log_{\lceil m/2 \rceil} n$ και $\log_m n$

Εξωτερική Αναζήτηση

B-δένδρα

Ένα (a,b) -δένδρο είναι δένδρο αναζήτησης πολλαπλής διακλάδωσης όπου:

- $a \geq 2$ και $b > a$ (συνήθως $b \geq 2 \cdot a$)
- η ρίζα έχει $d-1$ κλειδιά και d παιδιά, $2 \leq d \leq b$
- οι υπόλοιποι εσωτερικοί κόμβοι έχουν $t-1$ κλειδιά και t παιδιά, $a \leq t \leq b$
- οι κενοί κόμβοι (φύλλα) ισαπέχουν από τη ρίζα (δηλ. βρίσκονται στο ίδιο επίπεδο)

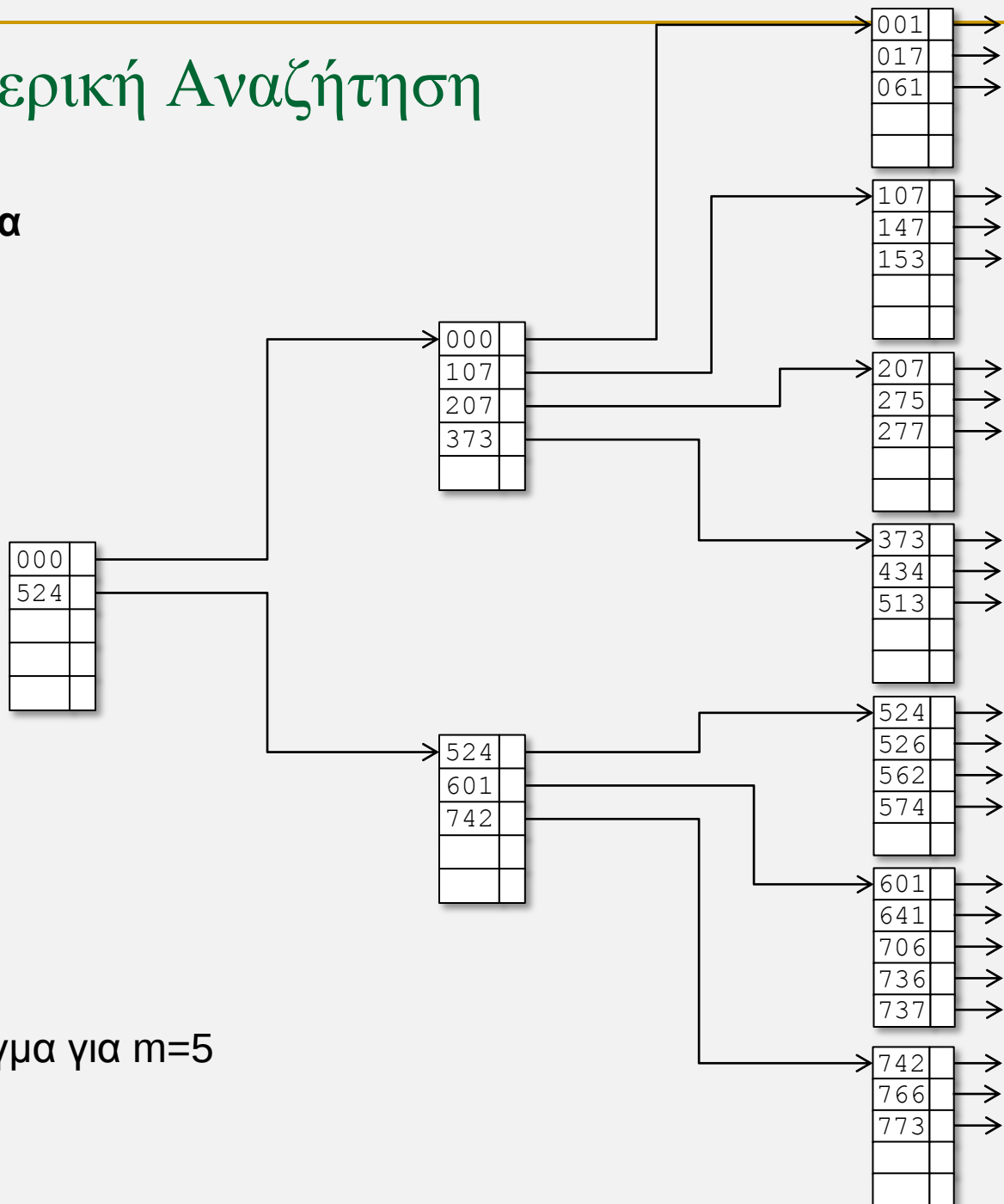
B-δένδρο (βαθμού m) : Αντιστοιχεί σε (a,b) -δένδρο με $a = \lceil m/2 \rceil, b = m$

Επιλέγουμε αρκετά μεγάλο m ώστε οι κόμβοι με m παιδιά να γεμίζουν μία σελίδα.

Επειδή για κάθε κλειδί πρέπει να αποθηκεύουμε δείκτη προς το αντίστοιχο στοιχείο αποθηκεύουμε τους δείκτες προς τα στοιχεία στους κόμβους του τελευταίου επιπέδου. Στους εσωτερικούς κόμβους διατηρούμε αντίγραφα των κλειδιών.

Εξωτερική Αναζήτηση

B-δένδρα



Παράδειγμα για $m=5$

Εξωτερική Αναζήτηση

Β-δένδρα Ακολουθία εισαγωγών με $m=5$

κλειδιά

706
176
601
153
→ 513
773
742
373
524
766
275
737
574
434
641
207
001
277
061
736
526
562
017
107
147

153		→
176		→
513		→
601		→
706		→

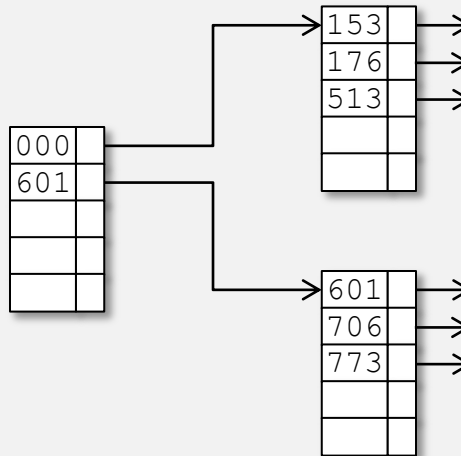
Εξωτερική Αναζήτηση

Β-δένδρα Ακολουθία εισαγωγών με $m=5$

κλειδιά

706
176
601
153
513
→ 773
742
373
524
766
275
737
574
434
641
207
001
277
061
736
526
562
017
107
147

ελάχιστο δυνατό κλειδί

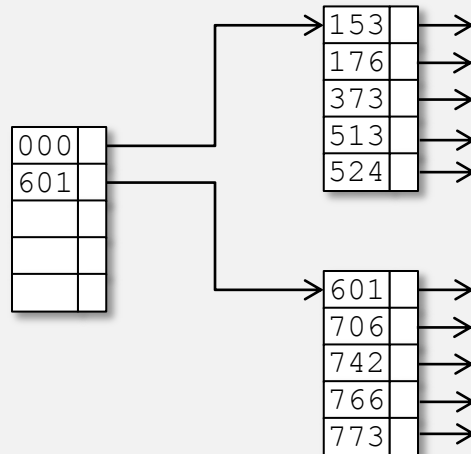


Εξωτερική Αναζήτηση

B-δένδρα Ακολουθία εισαγωγών με $m=5$

κλειδιά

706
176
601
153
513
773
742
373
524
→ 766
275
737
574
434
641
207
001
277
061
736
526
562
017
107
147



Εξωτερική Αναζήτηση

B-δένδρα Ακολουθία εισαγωγών με $m=5$

κλειδιά

706

176

601

153

513

773

742

373

524

766

→ 275

737

574

434

641

207

001

277

061

736

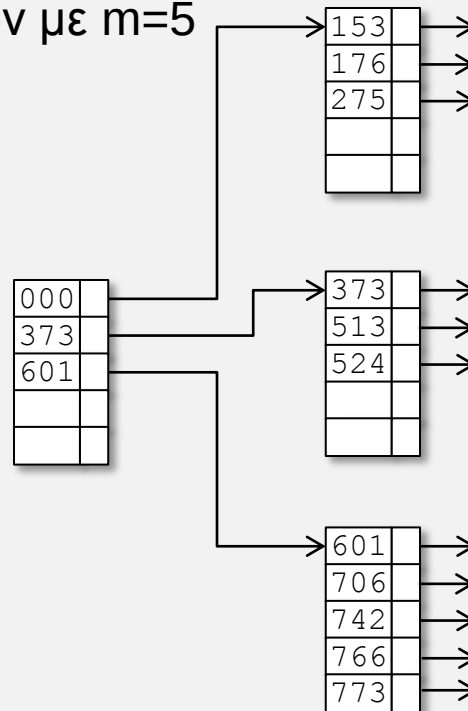
526

562

017

107

147

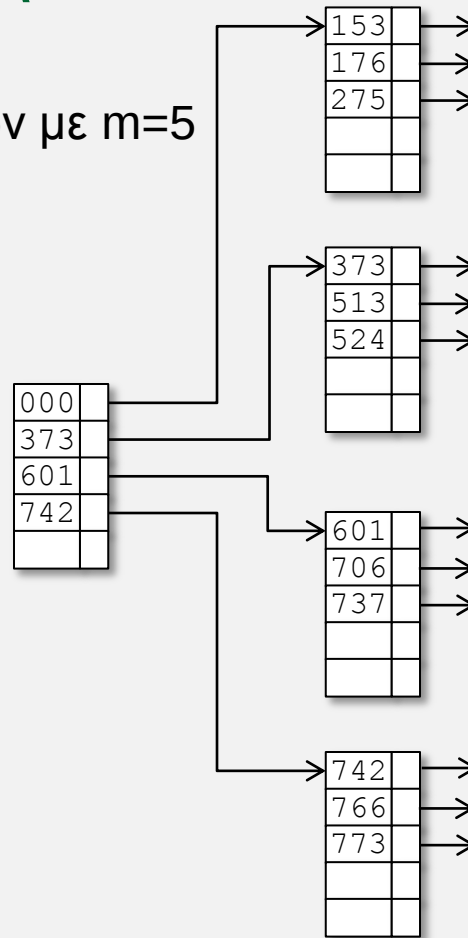


Εξωτερική Αναζήτηση

B-δένδρα Ακολουθία εισαγωγών με $m=5$

κλειδιά

706
176
601
153
513
773
742
373
524
766
275
→ 737
574
434
641
207
001
277
061
736
526
562
017
107
147

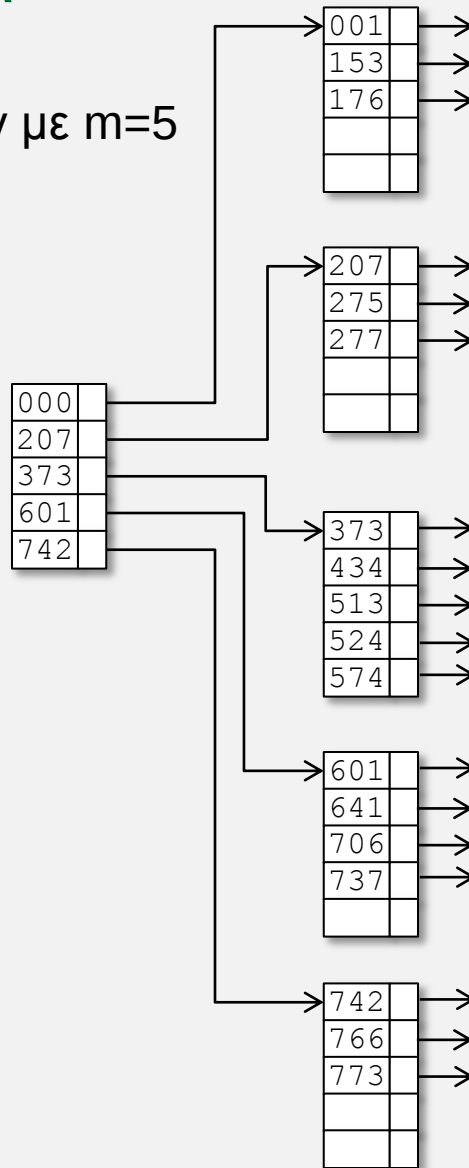


Εξωτερική Αναζήτηση

B-δένδρα Ακολουθία εισαγωγών με $m=5$

κλειδιά

706
176
601
153
513
773
742
373
524
766
275
737
574
434
641
207
001
→ 277
061
736
526
562
017
107
147

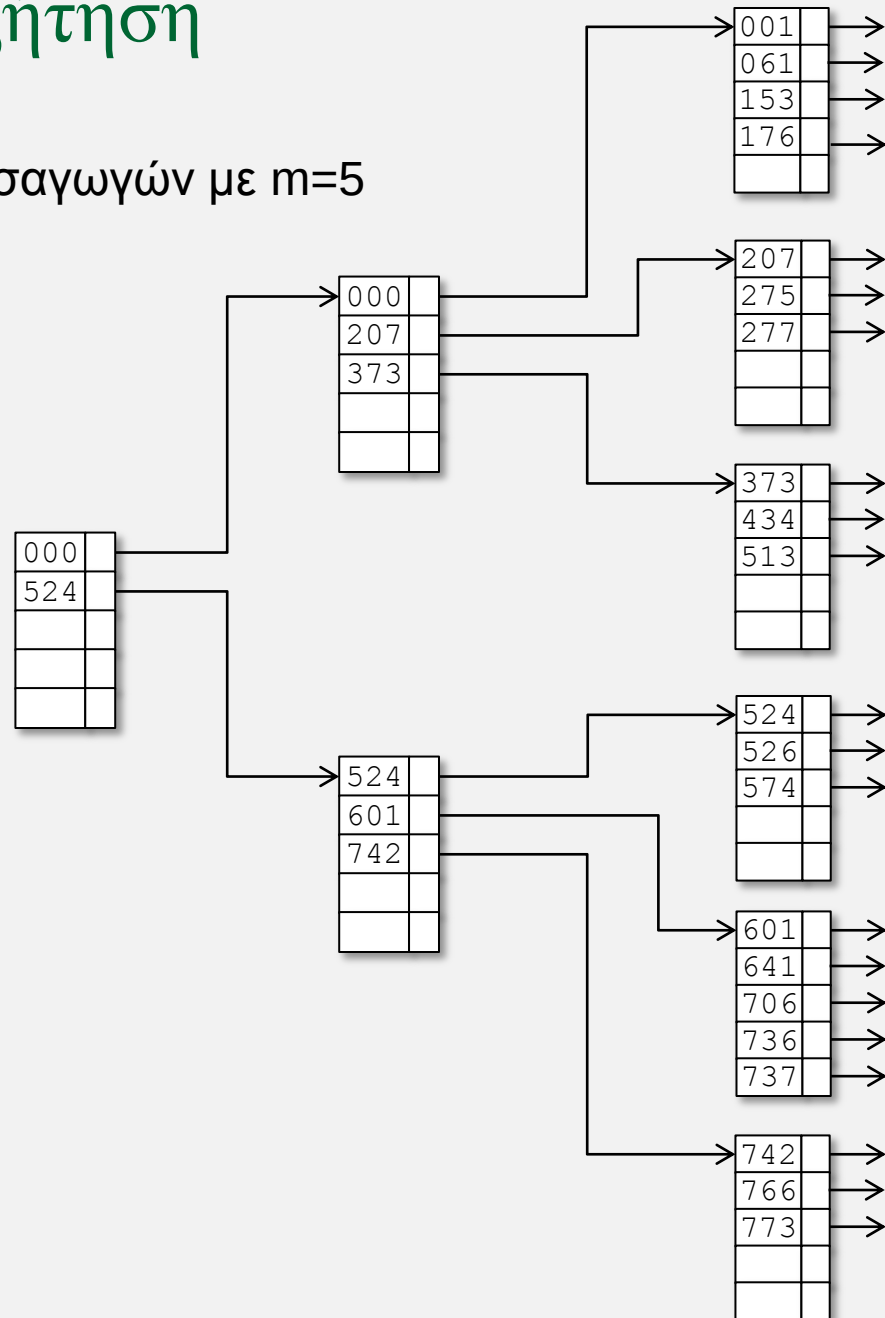


Εξωτερική Αναζήτηση

B-δένδρα Ακολουθία εισαγωγών με $m=5$

κλειδιά

706
176
601
153
513
773
742
373
524
766
275
737
574
434
641
207
001
277
061
736
→ 526
562
017
107
147

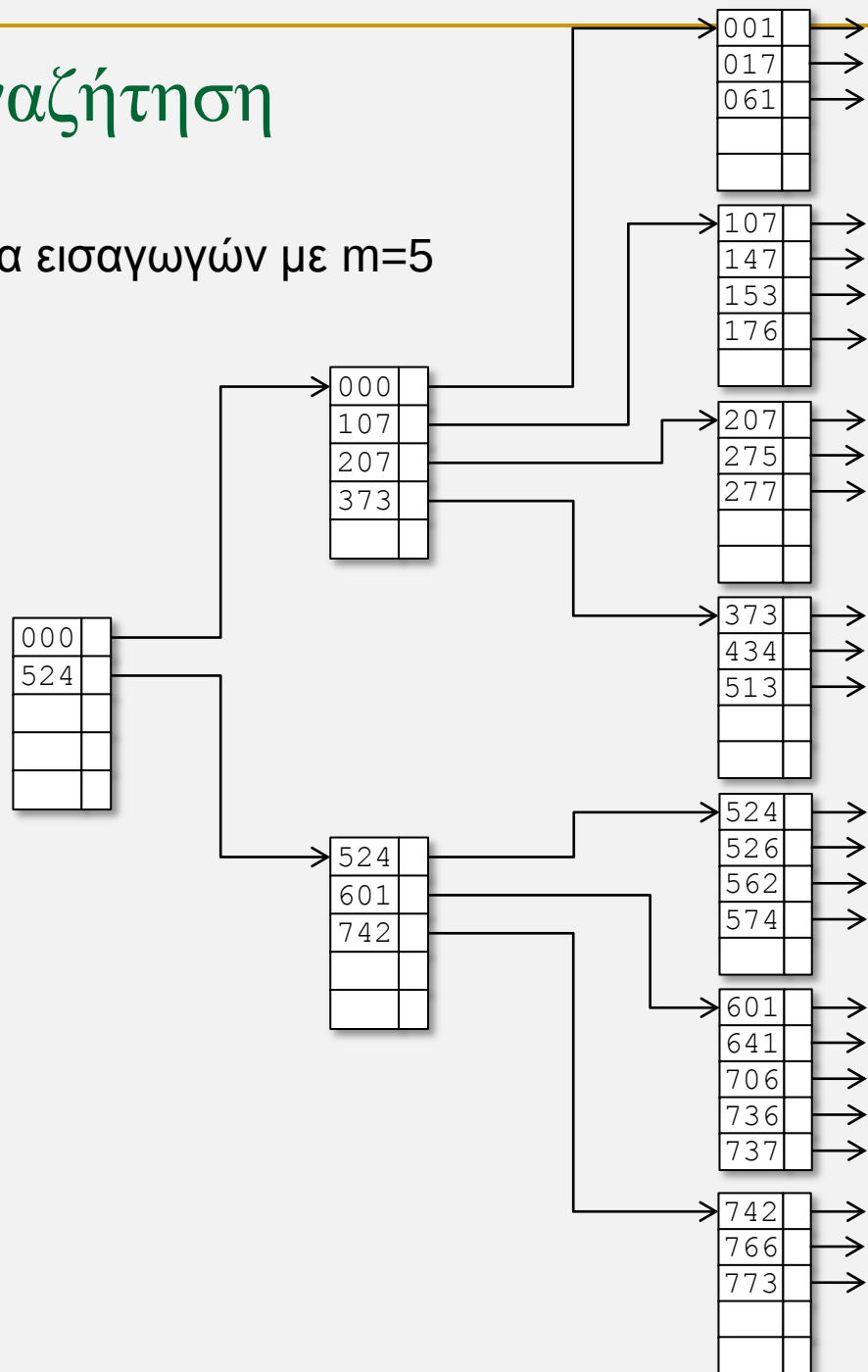


Εξωτερική Αναζήτηση

B-δένδρα Ακολουθία εισαγωγών με $m=5$

κλειδιά

706
176
601
153
513
773
742
373
524
766
275
737
574
434
641
207
001
277
061
736
526
562
017
107
→ 147



Εξωτερική Αναζήτηση

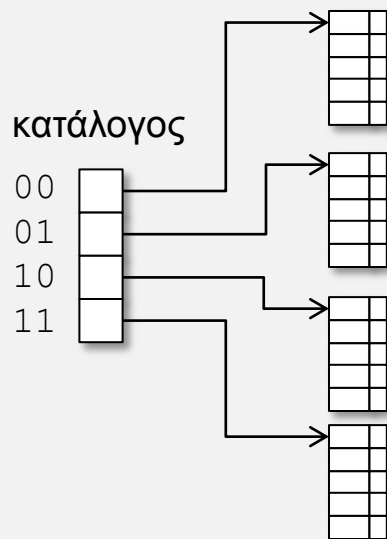
Επεκτάσιμος κατακερματισμός

Επιτρέπει την επαύξηση ή τη συρρίκνωση ενός πίνακα κατακερματισμού, διατηρώντας παράλληλα τους χρόνους πρόσβασης στη δομή χαμηλούς.

- Χρήσιμη μέθοδος για την αποθήκευση δεδομένων στη δευτερεύουσα μνήμη.
- Μπορεί να χρησιμοποιηθεί εναλλακτικά αντί ενός B-δένδρου.

Δομή 2 επιπέδων:

- Κατάλογος (πίνακας από δείκτες) που αποτελεί τη δομή υψηλού επιπέδου.
- Ένα σύνολο από σελίδες στις οποίες αποθηκεύονται (δείκτες προς) δεδομένα.

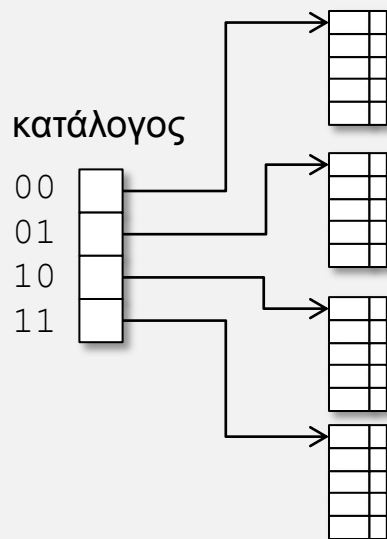


Εξωτερική Αναζήτηση

Επεκτάσιμος κατακερματισμός

Χρησιμοποιούμε μια συνάρτηση κατακερματισμού h για τη μετατροπή των κλειδιών σε ακέραιους.

Για την αναζήτηση ενός στοιχείου, πρώτα υπολογίζουμε την ακέραια τιμή της συνάρτησης κατακερματισμού και στη συνέχεια χρησιμοποιούμε τα πρώτα ψηφία της δυαδικής αναπαράστασης αυτού του ακεράιου για να εντοπίσουμε τη σελίδα που περιέχει το ζητούμενο στοιχείο.



Αν τα δεδομένα βρίσκονται σε N σελίδες, όπου

$$2^{D-1} < N \leq 2^D$$

τότε χρησιμοποιούμε τα πρώτα D ψηφία

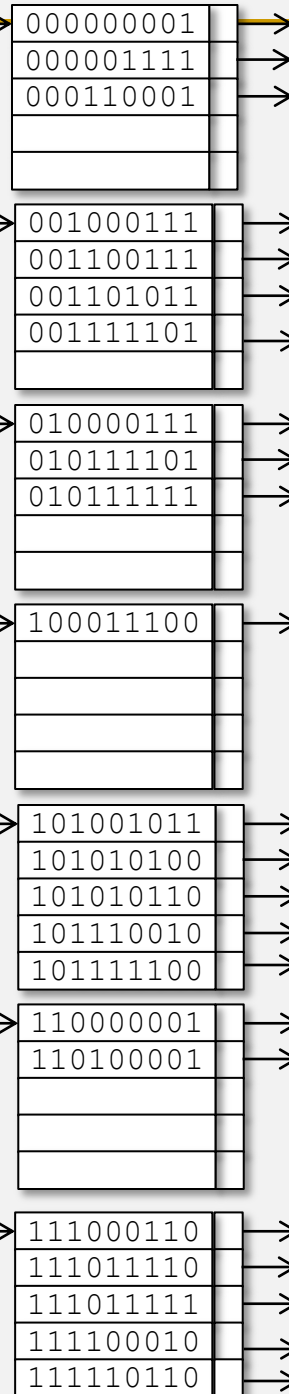
$$h_D(k) = \text{πρώτα } D \text{ δυαδικά ψηφία του } h(k)$$

Εξωτερική Αναζήτηση

τιμές συνάρτησης
κατακερματισμού
των κλειδιών

1110001110
0011111110
1100000001
0011010111
1010010111
1111110111
1111000010
0111110111
1010101000
1111101110
0101111101
1110111111
1011111100
1000111100
1101000001
0100000111
0000000001
0101111111
0001100001
1110111110
1010101110
1011100010
0000011111
0010000111
0011001111

000
001
010
011
100
101
110
111

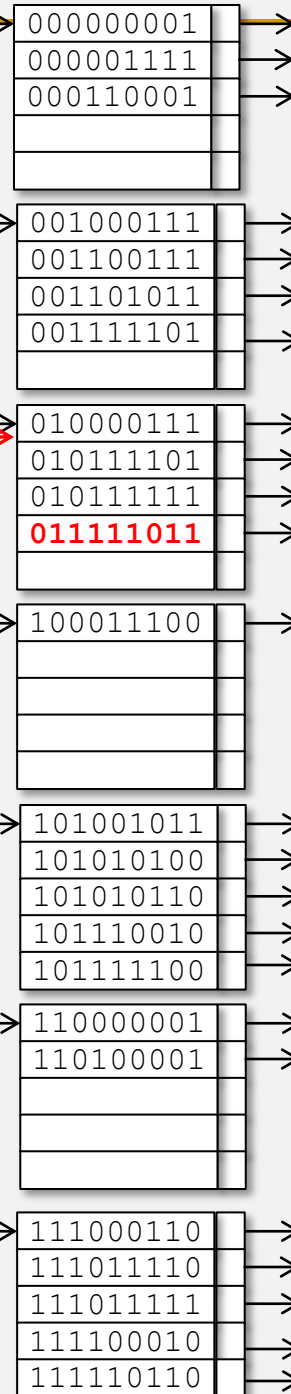


Εξωτερική Αναζήτηση

1110001110
0011111110
1100000001
0011010111
1010010111
1111111011
111100010
011111011
101010100
1111101110
010111101
1110111111
1011111100
1000111100
1101000001
010000111
0000000001
0101111111
000110001
1110111110
1010101110
101110010
000001111
001000111
001100111

τιμές συνάρτησης
κατακερματισμού
των κλειδιών

000
001
010
011
100
101
110
111



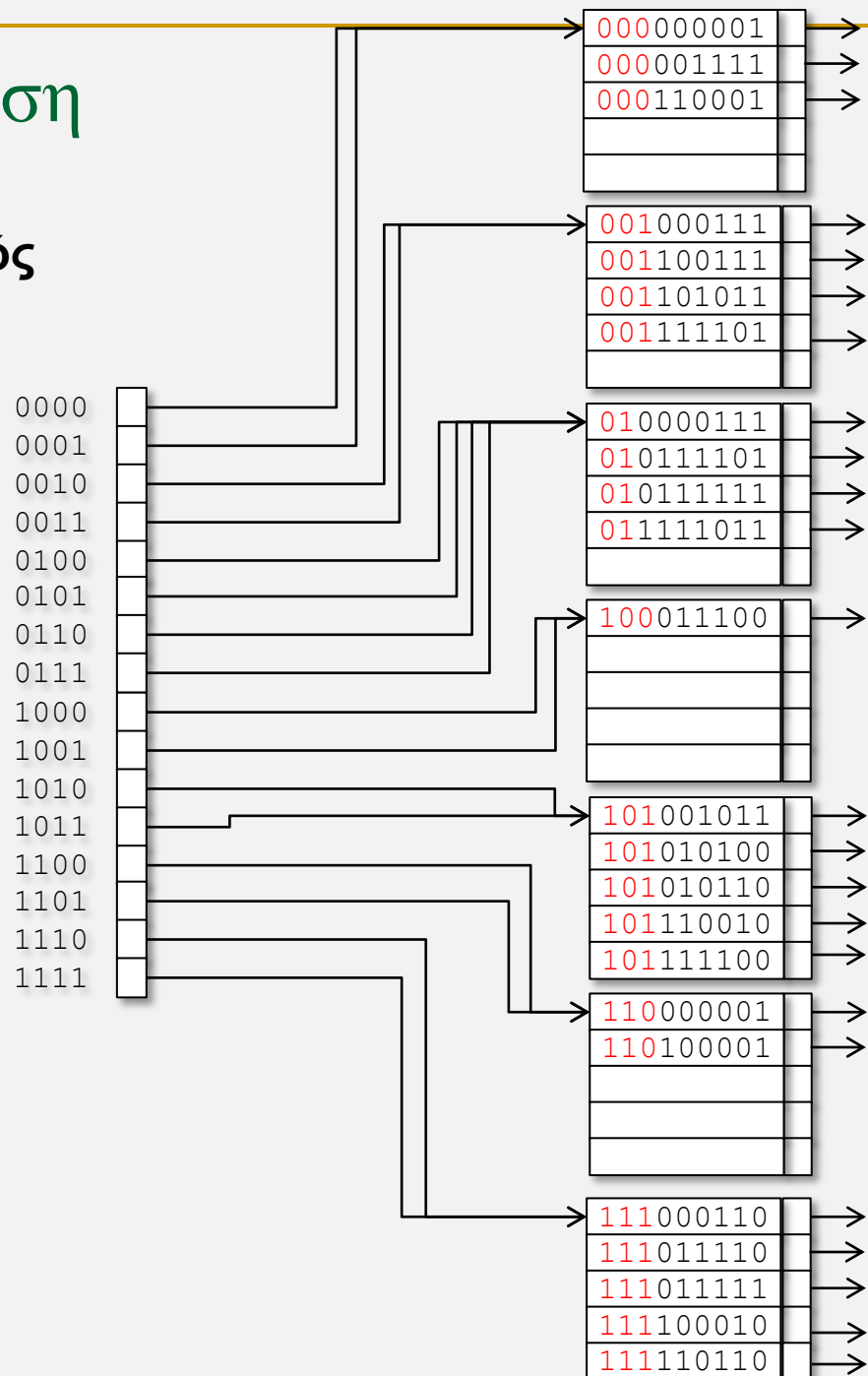
Αν κάποιες σελίδες δεν έχουν γεμίσει μπορούμε να μειώσουμε το πλήθος των σελίδων διατηρώντας πολλούς δείκτες του καταλόγου σε μια σελίδα

Εξωτερική Αναζήτηση

Επεκτάσιμος κατακερματισμός

Διπλασιασμός καταλόγου

Μπορεί να εξυπηρετήσει την
εισαγωγή περισσότερων
στοιχείων



Εξωτερική Αναζήτηση

Επεκτάσιμος κατακερματισμός

Μια σελίδα περιέχει όλα εκείνα τα κλειδιά των οποίων η τιμή κατακερματισμού έχει ένα συγκεκριμένο πρόθεμα από δυαδικά ψηφία. Το μήκος αυτού του προθέματος ονομάζεται **βάθος της σελίδας**. Το μέγιστο βάθος κάθε σελίδας ονομάζεται **βάθος του πίνακα κατακερματισμού**.

Ο κατάλογος είναι ένας πίνακας T με 2^D δείκτες σε σελίδες, όπου D το βάθος του πίνακα κατακερματισμού.

Εύρεση σελίδας που περιέχει το κλειδί k

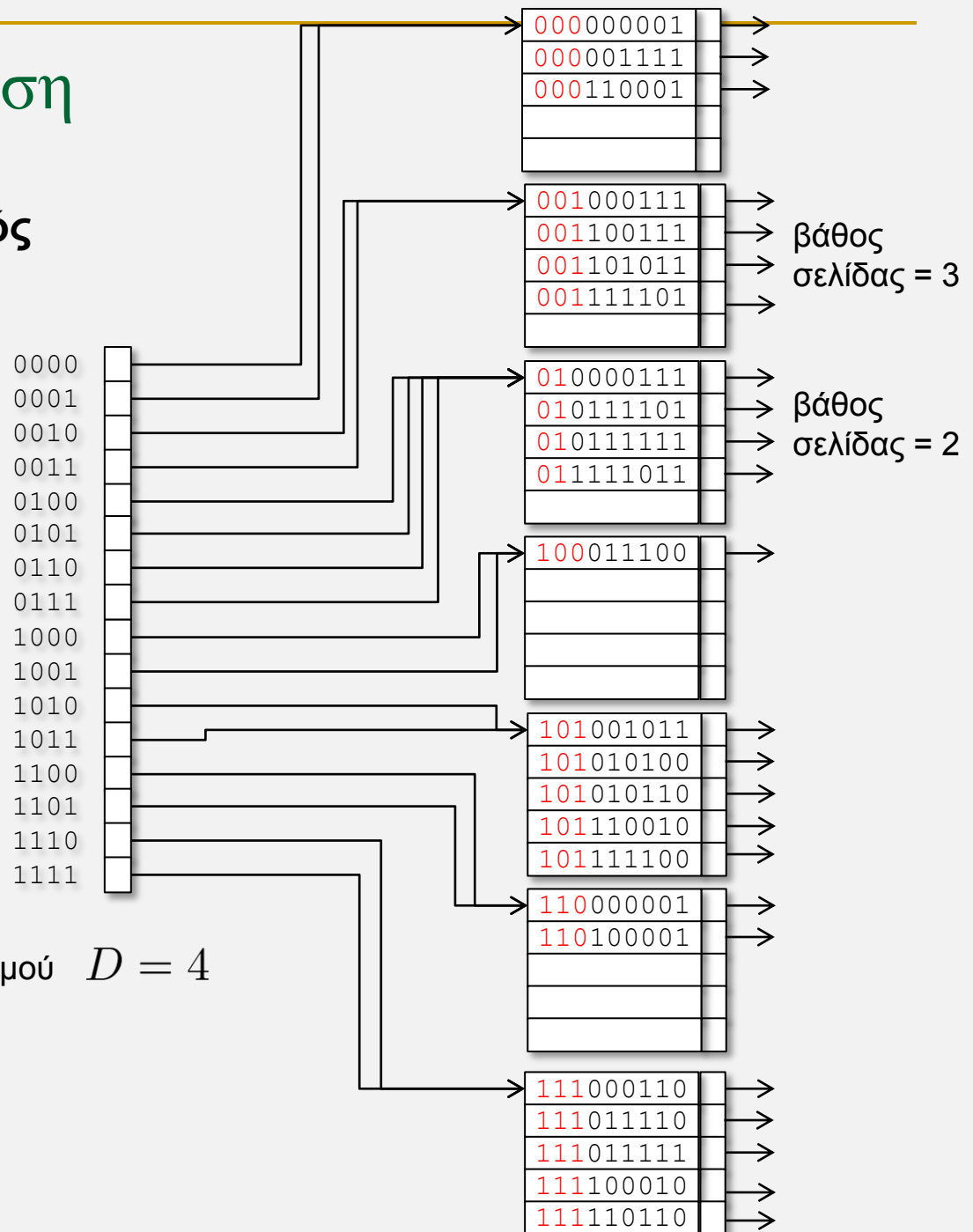
1. Υπολογίζουμε το $h_D(k)$
2. Ακολουθούμε τον δείκτη $T[h_D(k)]$

Αν μια σελίδα έχει βάθος $d < D$, τότε 2^{D-d} δείκτες σε συνεχόμενες θέσεις του πίνακα T θα δείχνουν στη σελίδα αυτή.

Εξωτερική Αναζήτηση

Επεκτάσιμος κατακερματισμός

βάθος πίνακα κατακερματισμού $D = 4$



Εξωτερική Αναζήτηση

Επεκτάσιμος κατακερματισμός

Εισαγωγές

Μια σελίδα χωράει μόνο m κλειδιά. Αν συμβεί **υπερχείλιση** μιας σελίδας με βάθος d , η σελίδα χωρίζεται σε δύο σελίδες:

- Δημιουργία μιας νέας σελίδας, που ονομάζεται **φιλική σελίδα (buddy page)** και
- Αύξηση του βάθους των σελίδων σε $d+1$.

Περιπτώσεις

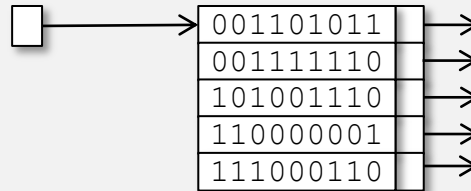
1. $d < D$: Αλλαγή απλά μερικών δεικτών στον κατάλογο, ώστε αυτοί να δείχνουν στη νέα σελίδα.
2. $d = D$: Διπλασιασμός του μεγέθους του καταλόγου και κατάλληλη αρχικοποίηση των δεικτών.

Εξωτερική Αναζήτηση

Επεκτάσιμος κατακερματισμός

Εισαγωγές

111000110
001111110
110000001
001101011
→ 101001011
111111011
111100010
011111011
101010100
111110110
010111101
111011111
101111100
100011100
010000111
000000001

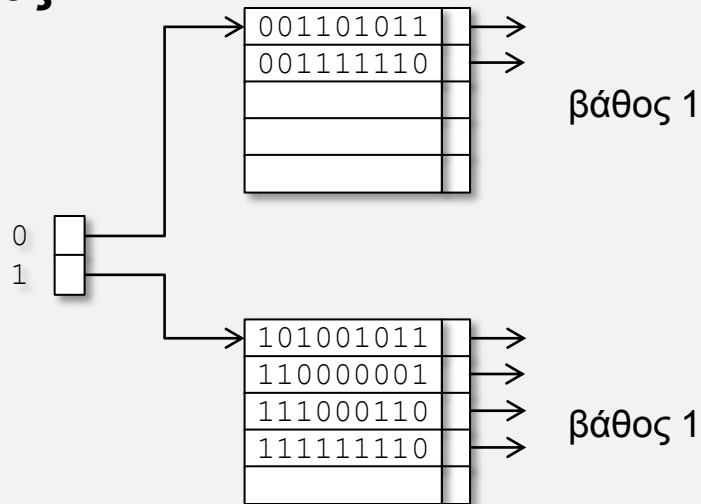


Εξωτερική Αναζήτηση

Επεκτάσιμος κατακερματισμός

Εισαγωγές

111000110
001111110
110000001
001101011
101001011
→ 111111011
111100010
011111011
101010100
111110110
010111101
111011111
101111100
100011100
010000111
000000001

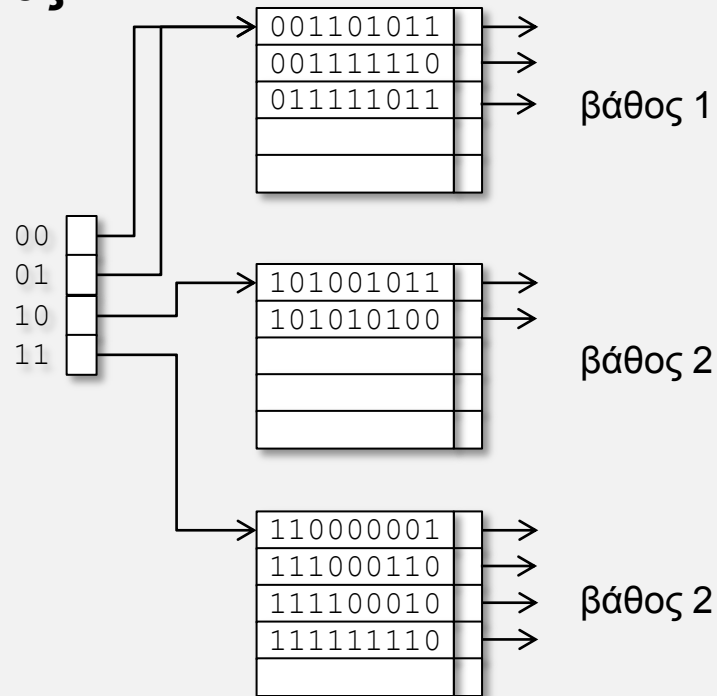


Εξωτερική Αναζήτηση

Επεκτάσιμος κατακερματισμός

Εισαγωγές

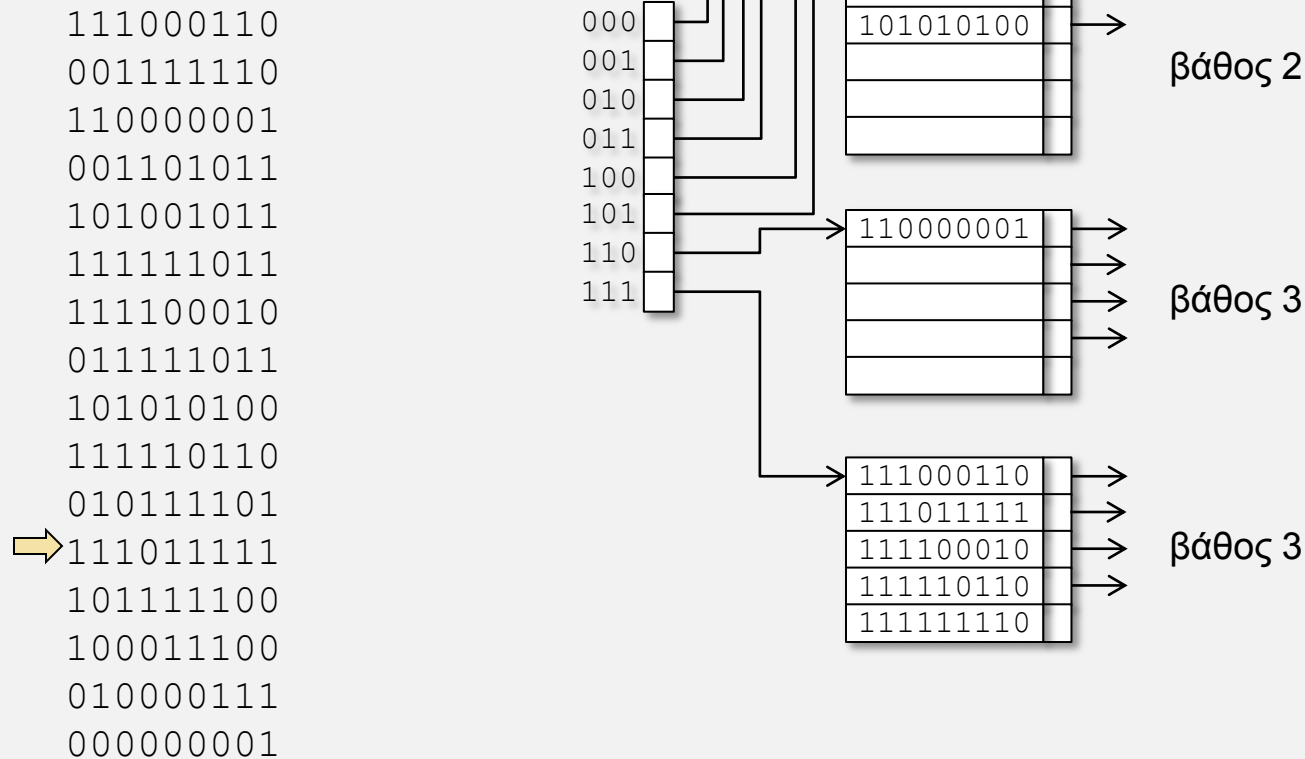
111000110
001111110
110000001
001101011
101001011
111111011
111100010
011111011
→ 101010100
111110110
010111101
111011111
101111100
100011100
010000111
000000001



Εξωτερική Αναζήτηση

Επεκτάσιμος κατακερματισμός

Εισαγωγές

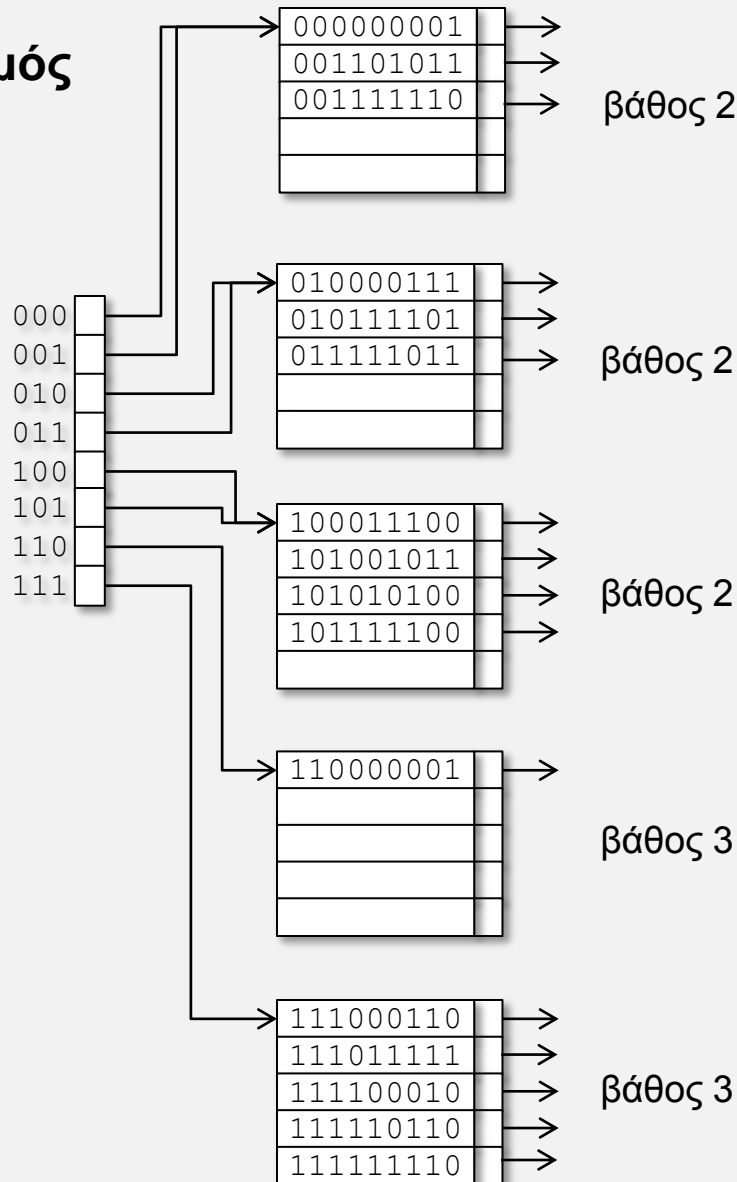


Εξωτερική Αναζήτηση

Επεκτάσιμος κατακερματισμός

Εισαγωγές

111000110
001111110
110000001
001101011
101001011
111111011
111100010
011111011
101010100
111110110
010111101
111011111
101111100
100011100
010000111
→ 000000001



Εξωτερική Αναζήτηση

Επεκτάσιμος κατακερματισμός

Διαγραφές

Αν η διαγραφή κάποιων κλειδιών από μια σελίδα $2i$, για κάποιο i , έχει ως αποτέλεσμα αυτή μαζί με την $(2i+1)$ να έχει συνολικά το πολύ m κλειδιά, θα μπορούσε να γίνει **συνένωση των 2 σελίδων σε 1**.

Γενικά αυτή η λειτουργία είναι **ακριβή και δεν είναι** συνετό να γίνεται συχνά.

Εξωτερική Αναζήτηση

Επεκτάσιμος κατακερματισμός

Αποδοτικότητα

(+) Είναι αποδοτικός σε χρόνο προσπέλασης : Απαιτείται μόνο μια πρόσβαση στην κύρια μνήμη και μια στην εξωτερική για να επιτευχθεί η προσπέλαση του ζητούμενου κλειδιού (και δεδομένου).

(-) Δεν είναι αρκετά αποδοτική σε μνήμη : Πολλές από τις σελίδες μπορεί να είναι σχεδόν άδειες και ο κατάλογος συνήθως περιέχει παραπάνω από έναν δείκτες που δείχνουν στην ίδια σελίδα.