

3^ο Σύνολο Ασκήσεων (προαιρετικό)

Καταληκτική Ημερομηνία Παράδοσης: Δευτέρα 14 Ιανουαρίου στις 12:00 στο γραφείο Β4
Θεματική Ενότητα: Αποθήκευση. Ευρετήρια. Εισαγωγή στην Επεξεργασία Ερωτήσεων.

Το σύνολο αυτό είναι ΠΡΟΑΙΡΕΤΙΚΟ και μετρά ΘΕΤΙΚΑ με την παρακάτω έννοια:
Ο βαθμός σας σε αυτό το σύνολο πολλαπλασιασμένος με 8% θα προστεθεί στον τελικό βαθμό σας στο μάθημα. Για παράδειγμα, αν ο τελικός βαθμός σας στο μάθημα είναι 6 και ο βαθμός στο 3^ο Σύνολο είναι 5, ο τελικός βαθμός σας θα αυξηθεί σε $6 + 0.08 \cdot 5 = 6.4$, ενώ αν ο βαθμός σας στο 3^ο Σύνολο είναι 10, ο τελικός βαθμός σας θα αυξηθεί σε $6 + 0.08 \cdot 10 = 6.8$.

Για γενικές οδηγίες σχετικά με τις ασκήσεις, συμβουλευτείτε την ιστοσελίδα του μαθήματος. Εξηγήστε τις απαντήσεις σας με σαφήνεια. Αν κάνετε οποιαδήποτε υπόθεση, καταγράψτε την.

Άσκηση 1

Θεωρείστε επεκτατό (extendible) κατακερματισμό με μέγεθος κάδου (bucket) 16 εγγραφών. Έστω ότι έχουμε 512 (2⁹) κλειδιά που κατακερματίζονται σε αριθμούς 12-bit. Όλα τα κλειδιά κατακερματίζονται σε διαφορετικούς αριθμούς.

(α) Δώστε ένα παράδειγμα αριθμών που να οδηγούν σε ένα κατάλογο με ολικό βάθος 5. Αν δεν υπάρχει, εξηγήστε γιατί δεν υπάρχει.

(β) Δώστε ένα παράδειγμα αριθμών που να οδηγούν σε ένα κατάλογο με ολικό βάθος 8. Αν δεν υπάρχει, εξηγήστε γιατί δεν υπάρχει.

Άσκηση 2

Θεωρείστε το ευρετήριο γραμμικού κατακερματισμού της Εικόνας 1.

(α) Εισάγετε στο ευρετήριο το 4 και το 15 και δώστε το αποτέλεσμα μετά από κάθε εισαγωγή.

(β) Εισάγετε στο αρχικό ευρετήριο της εικόνας το 2 το 15 και δώστε το αποτέλεσμα μετά από κάθε εισαγωγή.

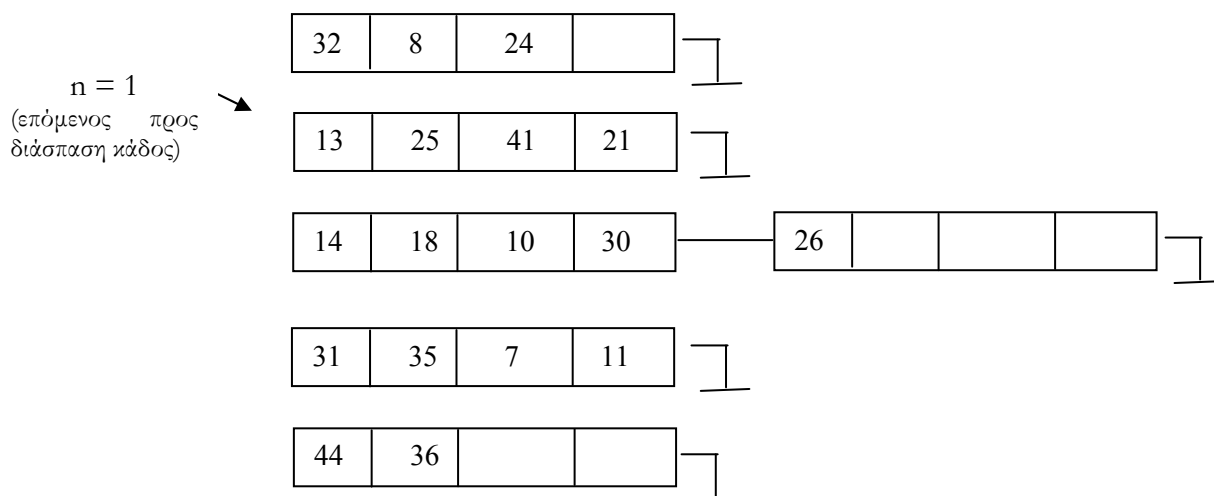
(γ) Ποια είναι η μικρότερη αριθμητική τιμή της οποίας η εισαγωγή στο ευρετήριο της Εικόνας 2 μπορεί να οδηγήσει στην υπερχείλιση ενός κάδου (δηλαδή, σε ένα κάδο με λίστα υπερχείλισης);

(δ) Στο ευρετήριο της Εικόνας 1 υπάρχουν 5 κάδοι (και 1 κάδος υπερχείλισης) και ο επόμενος προς διάσπαση κάδος είναι ο 1 (δηλαδή, ο δεύτερος κάδος). Υποθέστε ότι μετά από έναν αριθμό από εισαγωγές, υπάρχουν 25 κάδοι (χωρίς του κάδους υπερχείλισης). Ποιος θα είναι ο επόμενος κάδος προς διάσπαση;

$$h_0(k) = k \bmod 4$$

$$h_1(k) = k \bmod 8$$

Λίστες Υπερχείλισης



Εικόνα 1. Ευρετήριο για την Άσκηση 2. Κάθε κάδος χωρά 4 εγγραφές.

Άσκηση 3

Θεωρείστε ένα B+-δέντρο ύψους 4 (συμπεριλαμβανομένης της ρίζας), τάξης $p = 20$ και $p_{\text{leaf}} = 46$ σε ένα γνώρισμα A μιας σχέσης R . Ο παράγοντας ομαδοποίησης (σελιδοποίησης) του αρχείου της σχέσης είναι 30 και το A είναι κλειδί της R . Θεωρείστε τις δύο παρακάτω περιπτώσεις.

- (i) Το αρχείο είναι ταξινομημένο ως προς το A .
- (ii) Το αρχείο δεν είναι ταξινομημένο.

Απαντήστε στα παρακάτω ερωτήματα και για τις δύο περιπτώσεις.

- (α) Ποιο είναι το μεγαλύτερο αρχείο (σε αριθμό blocks) που μπορεί να δεικτοδοτήσει το ευρετήριο.
- (β) Περιγράψτε έναν αλγόριθμο για τον υπολογισμό της ερώτησης $\sigma_{a \leq A \leq b}(R)$ που να χρησιμοποιεί το ευρετήριο του ερωτήματος (α). Δώστε μια εκτίμηση της βελτίωσης του κόστους εκτέλεσης της ερώτησης (σε αριθμό blocks) όταν για τον υπολογισμό χρησιμοποιείται το ευρετήριο και όταν δε χρησιμοποιείται. Υποθέστε ότι την συνθήκη την ικανοποιεί το 5% των εγγραφών του αρχείου.
- (γ) Θεωρείστε την εισαγωγή μιας νέας τιμής στο αρχείο. Περιγράψτε το αλγόριθμο εισαγωγής και δώστε μια εκτίμηση του κόστους του (σε αριθμό blocks) στη χειρότερη περίπτωση.

Άσκηση 4

Θεωρήστε δύο σχέσεις R και S . Η R έχει 100.000 πλειάδες μεγέθους 100 bytes η καθεμία και η S έχει 20.000 μεγέθους 100 bytes η καθεμία. Κάθε καταχωρητής μνήμης (buffer) έχει μέγεθος 1 block. Το μέγεθος του block είναι 1024 bytes.

- (α) Θέλουμε να υπολογίσουμε το $R \cap S$ χρησιμοποιώντας έναν αλγόριθμο που διαβάζει κάθε block της R και της S μόνο μια φορά. Περιγράψτε τον αλγόριθμο που χρησιμοποιεί τον ελάχιστο αριθμό από καταχωρητές μνήμης. Ποιος είναι αυτός ο αριθμός; Οι R και S είναι αποθηκευμένες σε αρχεία σωρού.
- (β) Υποθέστε ότι η S είναι αποθηκευμένη σε διατεταγμένο αρχείο και το πεδίο διάταξης είναι και πρωτεύον κλειδί για την S . Επίσης, στην S υπάρχει ευρετήριο (ενός επίπεδο) πάνω στο πρωτεύον κλειδί. Το μέγεθος κάθε εγγραφής του ευρετηρίου είναι 10 bytes. Περιγράψτε έναν αλγόριθμο για τον υπολογισμό του $R \cap S$ που να χρησιμοποιεί το ευρετήριο στο S . Ποιο είναι το κόστος του; Θεωρήστε ότι έχετε μόνο δυο καταχωρητές μνήμης.