

3^ο Σύνολο Ασκήσεων

Καταληκτική Ημερομηνία Παράδοσης: Παρασκευή 14 Ιανουαρίου στις 15:00 στο γραφείο B15
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 15% 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 $6 + 0.15 \cdot 5 = 7$, ενώ αν ο βαθμός σε αυτό είναι 10, ο τελικός βαθμός σας θα αυξηθεί $7 + 0.15 \cdot 10 = 8.5$.

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

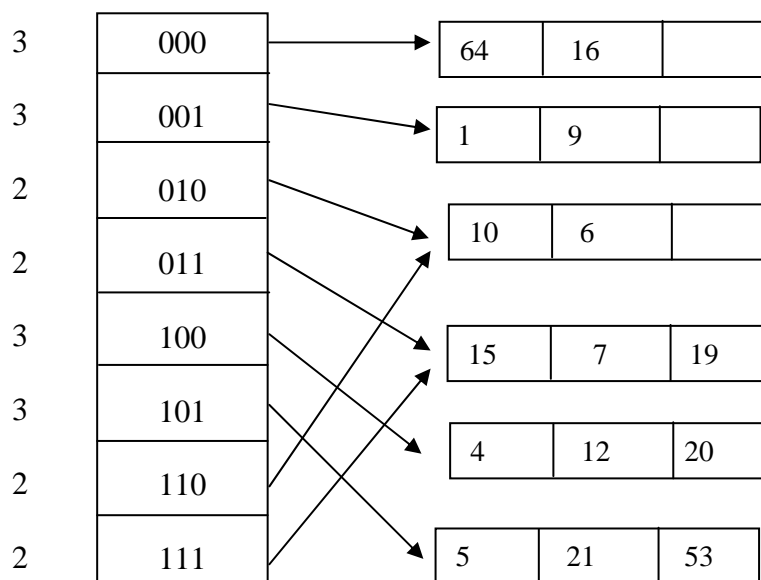
Άσκηση 1 [18 μονάδες]

Θεωρείστε το ευρετήριο επεκτατού κατακερματισμού της Εικόνας 1.

- (i) Ποιος είναι ο μεγαλύτερος αριθμός τιμών (δηλ. πλήθος τιμών) που μπορεί να εισαχθούν σε αυτό χωρίς να χρειαστεί να αυξηθεί το ολικό βάθος του καταλόγου; Εξηγήστε σε ποιους κάρδους πρέπει να πάνε αυτές οι τιμές.
- (ii) Ποιος είναι ο ελάχιστος αριθμός τιμών που μπορεί να εισαχθούν σε αυτό ώστε να χρειαστεί να αυξηθεί το ολικό βάθος του καταλόγου κατά 1; Δώστε ένα παράδειγμα τέτοιων τιμών.
- (iii) Ποιος είναι ο ελάχιστος αριθμός τιμών που μπορεί να εισαχθούν σε αυτό ώστε να χρειαστεί να αυξηθεί το ολικό βάθος του καταλόγου κατά 2; Δώστε ένα παράδειγμα τέτοιων τιμών.
- (iv) Στο μάθημα, μελετήσαμε αναλυτικά τις εισαγωγές τιμών και την αύξηση του μεγέθους του καταλόγου. Εξηγήστε πότε είναι δυνατή η μείωση του ολικού βάθους ενός καταλόγου. Ποιος είναι ο ελάχιστος αριθμός τιμών που η διαγραφή τους μπορεί να οδηγήσει σε μείωση κατά 1 του ολικού βάθους του καταλόγου της Εικόνας 1. Δώστε ένα παράδειγμα τέτοιων τιμών.
- (v) Το ευρετήριο της Εικόνας 1 έχει 6 κάρδους. Θεωρείστε ότι σε κάποια στιγμή το ευρετήριο φτάνει τους 200 κάρδους. Ποιο είναι το μικρότερο δυνατό ολικό βάθος σε αυτήν την περίπτωση; Ποιο είναι το μεγαλύτερο δυνατό ολικό βάθος σε αυτήν την περίπτωση;

Ολικό Βάθος = 3

Τοπικός Βάθος



Εικόνα 1. Ευρετήριο για την Άσκηση 1. Κάθε κάρδος χωρά 3 εγγραφές.

Άσκηση 2 [17 μονάδες]

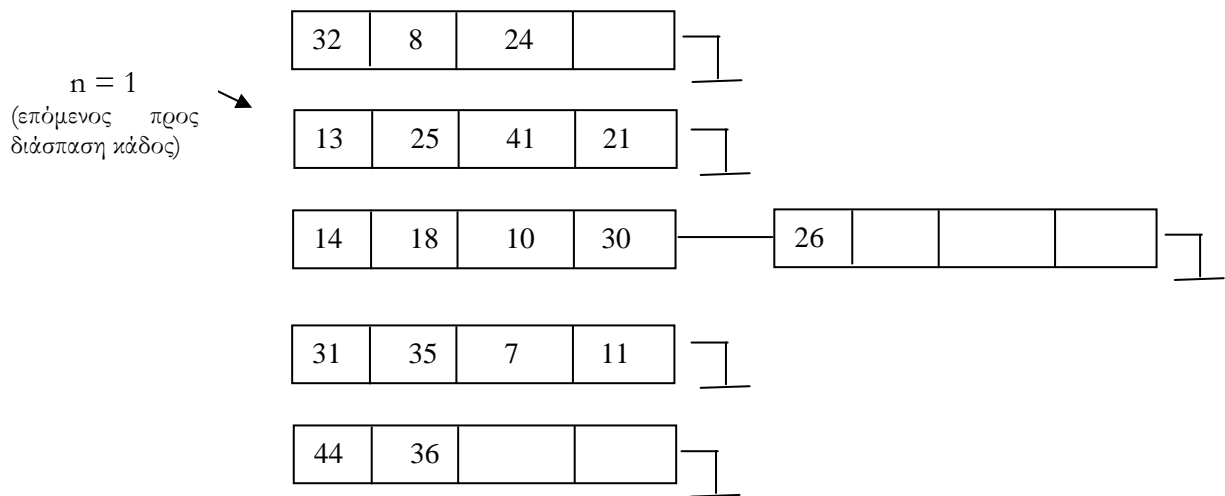
Θεωρείστε το ευρετήριο γραμμικού κατακερματισμού της Εικόνας 2.

- (i) Εισάγετε στο ευρετήριο το 4 και το 15 και δώστε το αποτέλεσμα μετά από κάθε εισαγωγή.
- (ii) Εισάγετε στο αρχικό ευρετήριο της Εικόνας 2 το 2 το 15 και δώστε το αποτέλεσμα μετά από κάθε εισαγωγή.
- (iii) Θεωρείστε το αρχικό ευρετήριο της Εικόνας 2 και ότι δεν έχουν προηγηθεί διαγραφές. Τι μπορείτε να πείτε για την τελευταία τιμή της οποίας η εισαγωγή προκάλεσε διάσπαση. Δώστε όλες τις πιθανές τέτοιες τιμές.
- (iv) Ποια είναι η μικρότερη αριθμητικά τιμή της οποίας η εισαγωγή στο ευρετήριο της Εικόνας 2 μπορεί να οδηγήσει στην υπερχείλιση ενός κάδου (δηλαδή, σε ένα κάδο με λίστα υπερχείλισης);
- (v) Στο ευρετήριο της Εικόνας, 2 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;

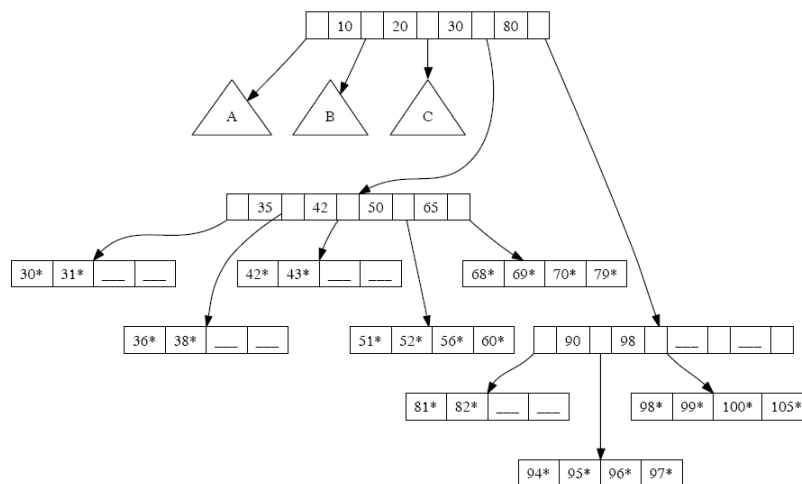
$$h_0(k) = k \bmod 4$$

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

Λίστες Υπερχείλισης



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



Εικόνα 3. B+ δέντρο για την Άσκηση 3. Η τάξη για τους εσωτερικούς κόμβους είναι 5 και για τα φύλλα 4. Τα φύλλα είναι διπλά δια-συνδεδεμένα μεταξύ τους (δε φαίνεται στην εικόνα). Τα A, B, C είναι υπο-δέντρα στα οποία δείχνουν οι αντίστοιχοι δείκτες.

Άσκηση 3 [15 μονάδες]

Θεωρείστε το B+ δέντρο της Εικόνας 3.

- (i) Σε ποιο υπο-δέντρο θα αναζητήσετε την τιμή 9 και σε ποιο την τιμή 25;
- (ii) Εισάγετε την τιμή 101 και δώστε το αποτέλεσμα της εισαγωγής.
- (iii) Διαγράψτε την τιμή 43 και δώστε το αποτέλεσμα της διαγραφής.
- (iv) Εισάγετε την τιμή 80 και δώστε το αποτέλεσμα της εισαγωγής.
- (v) Διαγράψτε την τιμή 30 και δώστε το αποτέλεσμα της διαγραφής.

Άσκηση 4 [50 μονάδες]

Θεωρείστε μια σχέση R που έχει 500.000 εγγραφές (πλειάδες) και είναι αποθηκευμένη σε ένα αρχείο σωρού.

Κατασκευάζουμε ένα ευρετήριο για ένα γνώρισμα (πεδίο) A της σχέσης R δεν είναι κλειδί για τη σχέση. Έστω ότι υπάρχουν 1.000 διαφορετικές τιμές του A. Για τα ερωτήματα (α), (β) και (γ) θεωρείστε ότι κατασκευάζουμε ένα ευρετήριο με *ενδιάμεσο επίπεδο*. Το μέγεθος του πεδίου A είναι 8 bytes και των δεικτών 16 bytes (όλων των ειδών δεικτών). Θεωρείστε μέγεθος block 2048 bytes.

(α) Έστω ότι το ευρετήριο είναι ένα δευτερεύον ευρετήριο (ενός επιπέδου).

(i) Δώστε το συνολικό μέγεθος του αρχείου ευρετηρίου (συμπεριλαμβανομένου και του ενδιάμεσου επιπέδου) σε αριθμό blocks.

(ii) Υπολογίστε το κόστος (σε αριθμό blocks) της αναζήτησης $A = \alpha$, όπου α μια τιμή από το πεδίο ορισμού του A. Υποθέστε ότι η αναζήτηση ικανοποιείται από το 0.1% των εγγραφών της σχέσης.

(iii) Έστω ότι εισάγουμε μια νέα πλειάδα στη σχέση R. Εξηγήστε πως αλλάζει το ευρετήριο, όταν η εγγραφή έχει μια τιμή στο A η οποία (1) δεν εμφανίζεται ήδη σε άλλη πλειάδα και (2) εμφανίζεται ήδη σε άλλη πλειάδα. Δώστε μια εκτίμηση του κόστους (πολυπλοκότητας) αυτών των αλλαγών (σε αριθμό blocks) για κάθε μία από τις δύο περιπτώσεις.

(β) Υποθέστε τώρα ότι κατασκευάζουμε ένα B+-δέντρο (οι δείκτες τιμών δείχνουν στο ενδιάμεσο επίπεδο). Επαναλάβετε το υπο-ερωτήματα (i)-(iii) του ερωτήματος (α) θεωρώντας ότι το δέντρο είναι όσο το δυνατόν πιο γεμάτο.

(γ) Υποθέστε τώρα ότι κατασκευάζουμε ένα B-δέντρο (οι δείκτες τιμών δείχνουν στο ενδιάμεσο επίπεδο). Επαναλάβετε το υπο-ερωτήματα (i) και (ii) του ερωτήματος (α) θεωρώντας ότι το δέντρο είναι όσο το δυνατόν πιο γεμάτο.

(δ) Υποθέστε τώρα ότι κατασκευάζουμε ένα *πυκνό* δευτερεύον ευρετήριο (δηλαδή, **δεν** υπάρχει ενδιάμεσο επίπεδο). Επαναλάβετε το υπο-ερωτήματα (i) και (ii) του ερωτήματος (α).

(ε) Υποθέστε μια αναζήτηση $A > \alpha$, όπου α μια τιμή από το πεδίο ορισμού του A, η οποία ικανοποιείται από το 10% των εγγραφών της σχέσης. Εξηγήστε πως μπορείτε να χρησιμοποιήσετε το ευρετήρια των ερωτημάτων (α)-

(δ) για αυτήν την αναζήτηση και δώστε μια εκτίμηση του κόστους της σε κάθε περίπτωση.