

### 3<sup>ο</sup> Σύνολο Ασκήσεων

Καταληκτική Ημερομηνία Παράδοσης: Πέμπτη 16 Ιανουαρίου 2014, στο μάθημα

Θεματική Ενότητα: Ευρετήρια

Για γενικές οδηγίες σχετικά με τις ασκήσεις, συμβουλευτείτε την ιστοσελίδα του μαθήματος.

**Άσκηση 1.** Έστω ο πίνακας `USER(USER_NAME, GENDER, NATIONALITY, DATE_OF_BIRTH)`. Υποθέστε ότι ο πίνακας έχει 100.000 πλειάδες και είναι αποθηκευμένος σε ένα αρχείο ταξινομημένο με πεδίο ταξινόμησης το γνώρισμα `USER_NAME`. Το γνώρισμα `USER_NAME` έχει μέγεθος 8 bytes, τα γνωρίσματα `GENDER` και `NATIONALITY` από 1 byte και το γνώρισμα `DATE_OF_BIRTH` 3 bytes. Ο δείκτης αρχείου έχει μέγεθος 5 bytes. Θεωρείστε μέγεθος σελίδας (block) ίσο με 512 bytes.

(α) Υπολογίστε πόσες σελίδες καταλαμβάνει το αρχείο.

(β) Θεωρείστε ότι υπάρχει ένα πρωτεύον B+-δέντρο ευρετήριο με πεδίο ευρετηριοποίησης το γνώρισμα `USER_NAME`. Υπολογίστε το μικρότερο και μεγαλύτερο αριθμό σελίδων που θα καταλαμβάνει αυτό το ευρετήριο, υποθέτοντας ότι το μέγεθος δείκτη δέντρου είναι 4 bytes. Ποιο είναι το ύψος του δέντρου σε κάθε περίπτωση;

(γ) Θεωρείστε ότι υπάρχει ένα (στατικό) ευρετήριο κατακερματισμού με πεδίο ευρετηριοποίησης το γνώρισμα `USER_NAME`, όπου κάθε κάδος έχει μέγεθος ίσο με μία σελίδα. Πόσους κάδους θα έχει αυτό το ευρετήριο αν υποθέσουμε ότι κάθε κάδος είναι γεμάτος κατά 60% και οι εγγραφές του είναι ομοιόμορφα καταναμημένες;

(δ) Θεωρείστε την αναζήτηση εγγραφών που ικανοποιούν τη συνθήκη  $u1 < USER\_NAME < u2$ , όπου  $u1$  και  $u2$  είναι τιμές του πεδίου τιμών του γνωρίσματος `USER_NAME`. Έστω ότι την συνθήκη ικανοποιούν 1.000 πλειάδες. Ποιο είναι το κόστος (δηλαδή, αριθμός σελίδων που πρέπει να διαβαστούν) αν (i) δεν χρησιμοποιηθεί κανένα ευρετήριο, (ii) χρησιμοποιηθεί το μικρότερο B+-δέντρο και (iii) χρησιμοποιηθεί το ευρετήριο κατακερματισμού;

#### Άσκηση 2.

Επανάλαβατε τα ερωτήματα (β), (γ) και (δ) της Άσκησης 1 για την περίπτωση που ο πίνακας `USER` είναι αποθηκευμένος σε ένα αρχείο σωρού. Για το ερώτημα (β) υπολογίστε το μέγεθος και το ύψος μόνο του μικρότερου δέντρου.

#### Άσκηση 3.

Θεωρείστε επεκτατό (extendible) κατακερματισμό και υποθέστε ότι κάθε κάδος χωρά έως 3 εγγραφές. Εισάγετε τις παρακάτω τιμές σε ένα αρχικά άδειο κατάλογο. Σε παρένθεση δίνετε η δυαδική αναπαράσταση του αποτελέσματος του κατακερματισμού.

- a (001100)
- b (001100)
- c (101101)
- d (010010)
- e (111111)
- f (010010)
- g (100001)
- h (001100)
- i (000000)

(α) Δώστε τον κατάλογο μόνο μετά την εισαγωγή των τιμών που οδηγεί σε αύξηση του ολικού βάρους καθώς και το τοπικό βάρος κάθε εγγραφής του.

(β) Πόσους κάδους έχει ο τελικός κατάλογος;

(γ) Ποιος είναι ο ελάχιστος αριθμός τιμών που θα οδηγήσει σε αύξηση κατά 1 του ολικού βάρους (δηλαδή, διπλασιασμό) του καταλόγου; Δώστε ένα παράδειγμα αυτών των τιμών (τη δυαδική τους αναπαράσταση).

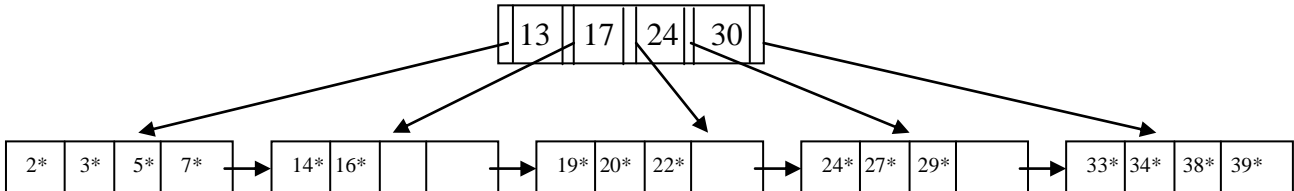
(δ) Ποιος είναι ο ελάχιστος αριθμός τιμών που θα οδηγήσει σε αύξηση κατά 2 του ολικού βάρους; Δώστε ένα παράδειγμα αυτών των τιμών (τη δυαδική τους αναπαράσταση).

#### Άσκηση 4

Θεωρείστε γραμμικό κατακερματισμό με  $M = 2$  και 3 εγγραφές ανά κάδο. Εισάγετε σε μια αρχικά άδεια δομή τις τιμές 0, 12, 33, 18, 63, 20, 7, 44, 24, 13, 32 με αυτήν τη σειρά. Θεωρείστε (όπως στο μάθημα) ότι γίνετε διάσπαση κάθε φορά που μια τιμή εισάγεται σε ένα γεμάτο κάδο (δηλαδή, όταν γίνεται υπερχείλιση). Δώστε τη δομή μετά την εισαγωγή κάθε μίας από τις παραπάνω τιμές.

#### Άσκηση 5

Έστω το παρακάτω B+ δέντρο.



- (α) Δώστε το δέντρο που προκύπτει μετά την εισαγωγή της τιμής 4.
- (β) Στο μάθημα περιγράψαμε μια τεχνική ανακατανομής εγγραφών με αδελφούς κόμβους σε περίπτωση υποχείλισης ενός φύλλου κατά τη διαγραφή τιμών. Μια παρόμοια ανακατανομή θα μπορούσε να χρησιμοποιηθεί και σε περίπτωση υπερχείλισης κατά την εισαγωγή τιμών. Εξηγήστε και δώστε το δέντρο που θα προέκυπτε μετά την εισαγωγή του 4 σε αυτήν την περίπτωση.
- (γ) Δώστε το B-δέντρο τάξης 3 που δεικτοδοτεί τις ίδιες τιμές.