

3^ο Σύνολο Ασκήσεων

Καταληκτική Ημερομηνία Παράδοσης: Τρίτη 9 Ιανουαρίου 2017

Το σύνολο θα λυθεί στη διάλεξη της Πέμπτης 11 Ιανουαρίου – κατά συνέπεια δε θα γίνουν δεκτές αργοπορημένες παραδόσεις ασκήσεων μετά την διάλεξη

Θεματική Ενότητα: Αποθήκευση, Ευρετήρια, Εισαγωγή στην επεξεργασία ερωτήσεων.

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

Άσκηση 1

Θεωρείστε ένα ευρετήριο επεκτατού κατακερματισμού, όπου κάθε κάδος (bucket/block) μπορεί να χωρέσει έως 3 εγγραφές.

(i) Θεωρείστε ότι κάποια στιγμή ο κατάλογος του ευρετηρίου έχει ολικό βάθος 3. Ποιο είναι το μικρότερο και ποιο το μεγαλύτερο δυνατό τοπικό βάθος; Δώστε ένα παράδειγμα τιμών των οποίων η εισαγωγή οδηγεί σε ένα ευρετήριο όπου κάποιες θέσεις έχουν αυτό το μικρότερο δυνατό τοπικό βάθος.

(ii) Θεωρείστε ότι κάποια στιγμή το ευρετήριο έχει 200 κάδους. Ποιο είναι το μικρότερο δυνατό ολικό βάθος για αυτόν τον κατάλογο;

Άσκηση 2

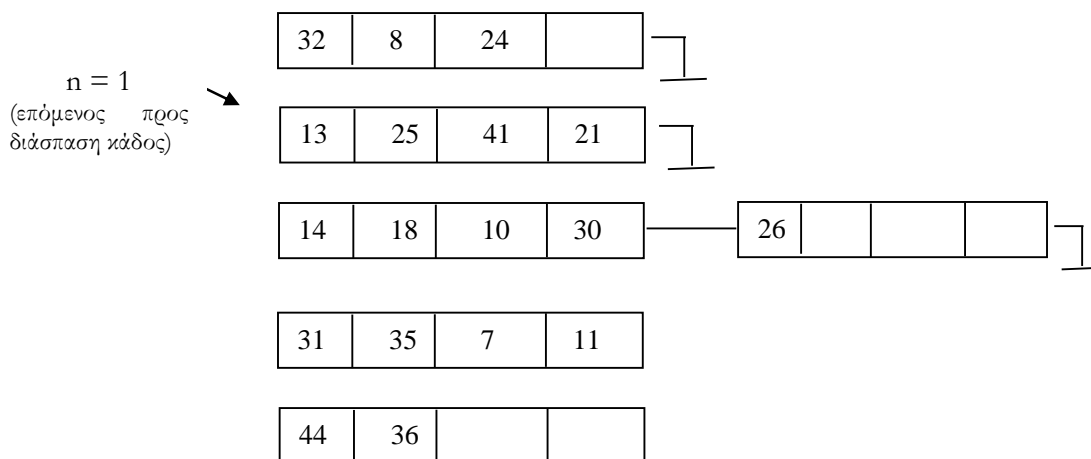
Θεωρείστε το ευρετήριο γραμμικού κατακερματισμού της Εικόνας 1, όπου υπάρχουν 5 κάδοι (και 1 κάδος υπερχείλισης) και ο επόμενος προς διάσπαση κάδος είναι ο 1 (δηλαδή, ο δεύτερος κάδος). Κάθε κάδος χωρά 4 εγγραφές.

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

(β) Εισάγετε στο αρχικό ευρετήριο της Εικόνας 1 το 2 και μετά το 15 και δώστε το αποτέλεσμα μετά από κάθε εισαγωγή.

(γ) Ποια είναι η μικρότερη αριθμητικά τιμή της οποίας η εισαγωγή στο ευρετήριο της Εικόνας 1 μπορεί να οδηγήσει στην υπερχείλιση ενός κάδου (δηλαδή, σε ένα κάδο με λίστα υπερχείλισης);

(δ) Υποθέστε ότι μετά από έναν αριθμό από εισαγωγές, υπάρχουν 25 κάδοι (χωρίς τους κάδους υπερχείλισης). Ποιος θα είναι ο επόμενος κάδος προς διάσπαση;



Εικόνα 1: Γραμμικό ευρετήριο για την Άσκηση 2

Άσκηση 3

Έστω ένας πίνακας (σχέση) CITY(Name, Population, Country) ο οποίος έχει πληροφορία για 100.000 πόλεις που είναι ομοιόμορφα κατανεμημένες σε 1.000 χώρες (δηλαδή 100 πόλεις ανά χώρα). Ο πίνακας είναι αποθηκευμένος σε ένα διατεταγμένο αρχείο ως προς το γνώρισμα Name. Τα γνωρίσματα Name και Country έχουν μέγεθος 16 bytes, το γνώρισμα Population 32 bytes και ένα block (σελίδα) 2048 bytes. Υποθέστε ότι όλοι οι δείκτες έχουν μέγεθος 16 bytes.

Θεωρείστε ότι υπάρχει ένα (μη δεντρικό) ευρετήριο στο γνώρισμα Country και ένα πυκνό B+-δέντρο στο γνώρισμα Population.

Έστω η ερώτηση

```
SELECT * FROM CITY WHERE Country = 'Greece' AND Population BETWEEN 50000 AND 60000;
```

Θεωρείστε τους παρακάτω διαφορετικούς τρόπους για να υπολογιστεί η ερώτηση

(α) με χρήση του B+-ευρετήριο στο γνώρισμα Population

(β) με χρήση του ευρετηρίου στο γνώρισμα Country

(γ) με χρήση και των δύο ευρετηρίων

(δ) χωρίς χρήση ευρετηρίου.

Υπολογίστε το κόστος καθενός από τα παραπάνω. Υποθέστε ότι υπάρχουν ότι υπάρχουν συνολικά 35 πόλεις με πληθυσμό πάνω από 50.000 έως 60.000 κατοίκους από τις οποίες ελληνικές είναι οι 10.