Η/Υ 4-32: Δομές Δεδομένων
Φθινόπωρο 2001 – Παναγιώτα Φατούρου

4ο Σετ Ασκήσεων

Ημερομηνία Παράδοσης
Θεωρητικό Μέρος: Τετάρτη 9/1/02, ώρα 13:00
Προγραμματιστικό Μέρος: Τετάρτη 9/1/02, ώρα 12:59 π.μ.
Τρόπος Παράδοσης: Προγραμματιστικό Μέρος: Ίδιος με εκείνον του 1ου Σετ Ασκήσεων.
Θεωρητικό Μέρος: Θα βάλετε τις ασκήσεις σας σε ένα φάκελο, θα γράψετε επάνω το όνομα σας και θα ρίξετε το φάκελο είτε στη θυρίδα μου ή κάτω από την πόρτα του γραφείου μου.

Θεωρητικό Μέρος

1. α. Ποιες είναι οι ιδιότητες μιας καλής συνάρτησης κατακερματισμού;

   β. Ποια είναι τα πλεονεκτήματα και μειονεκτήματα του κατακερματισμού σε σχέση με τις άλλες μεθόδους αναζήτησης;

   γ. Να καταταχθούν οι μέθοδοι κατακερματισμού που διδαχθήκατε ως προς της αλγοριθμική ευκολία τους και την πολυπλοκότητά τους κατά την αναζήτηση, την εισαγωγή και τη διαγραφή.

   δ. Σε πίνακα κατακερματισμού 300 θέσεων εισάγονται 30 τυχαία κλειδιά. Ποια είναι η πιθανότητα να μην υπάρχουν συγκρούσεις; [10%]

2. Χρησιμοποιώντας τη συνάρτηση κατακερματισμού της διαίρεσης ως πρωτεύουσα συνάρτηση κατακερματισμού και τη συνάρτηση $h(K) = K^3$ % m ως δευτερεύουσα συνάρτηση κατακερματισμού να εισαχθούν τα κλειδιά 73, 15, 44, 37, 30, 59, 49, 99 και 47 σε ένα πίνακα 10 θέσεων, ο οποίος είναι οργανωμένος με τη μέθοδο:
   a. της αλυσίδας
   b. των μικτών αλυσίδων
   c. της γραμμικής αναζήτησης
   d. του διπλού κατακερματισμού (όπου η συνάρτηση διαίρεσης παίζει το ρόλο και της δευτερεύουσας συνάρτησης)
   e. του ταξινομημένου κατακερματισμού με αλυσίδες
   f. του ταξινομημένου κατακερματισμού με ανοικτή διεύθυνση [10%]

3. α. Έστω ότι τα 3 κλειδιά με τιμές κατακερματισμού 01101, 01100 και 01000 εισάγονται στον πίνακα κατακερματισμού της διαφάνειας 21 της Ενότητας 7. Ζητούνται οι πίνακες κατακερματισμού που προκύπτουν μετά από κάθε εισαγωγή.

   β. Υποθέστε ότι χρησιμοποιείται μια στρατηγική διαγραφής σύμφωνα με την οποία θα πρέπει το directory (ο κατάλογος) να διατηρείται όσο το δυνατό μικρότερο. Ζητείται ο πίνακας κατακερματισμού που προκύπτει αν διαγραφεί το κλειδί με τιμή κατακερματισμού 01110 από τον επεκτάσιμο πίνακα κατακερματισμού της

διαφάνειας 21 της Ενότητας 7. Τι θα συμβεί αν διαγραφούν και τα κλειδιά με τιμές κατακερματισμού 01001 και 01010 από τον πίνακα;

γ. Ζητούνται οι πίνακες κατακερματισμού που προκύπτουν μετά την εισαγωγή κάθε ενός από τα κλειδιά με τιμές κατακερματισμού 000, 001, 010, 011, 100, 101, 110 και 111 σε έναν αρχικά άδειο επεκτάσιμο πίνακα κατακερματισμού. [10%]

4. α. Ζητούνται οι σοροί που προκύπτουν από την εισαγωγή των 6,5,2,4,7,1 και 3 σε έναν αρχικά άδειο σορό.

   β. Ζητούνται οι σοροί που προκύπτουν μετά από κάθε μια από 3 διαδοχικές εκτελέσεις της DeleteMin() στο σορό που προκύπτει στο ερώτημα α. [10%]

5. Δείξτε ότι η διαγραφή ενός οποιουδήποτε στοιχείου ενός σορού μπορεί να πραγματοποιηθεί σε χρόνο O(log n), αν έχουμε έναν δείκτη στον κόμβο (δηλαδή, σας ζητείται μια O(log n) υλοποίηση της HeapDelete(j,h), η οποία διαγράφει το στοιχείο με δείκτη j από τον σορό). [10%]

6. Υποθέστε ότι H_1 και H_2 είναι δύο σοροί μεγέθους n_1 και n_2, $n_1 \geq n_2$, ενώ κάθε στοιχείο του H_2 είναι μεγαλύτερο από κάθε στοιχείο του H_1. Εξηγήστε πως μπορούμε να συνενώσουμε τους δύο σορούς σε έναν σε χρόνο $O(n_1)$. [15%]

7. Οι σοροί επιτρέπουν την εύρεση του ελάχιστου στοιχείου σε χρόνο O(log n). Εξηγήστε πως η δομή του σορού θα μπορούσε να τροποποιηθεί έτσι ώστε να αποτελεί μια διπλή σορό, δηλαδή μια δομή δεδομένων που υποστηρίζει τις εξής λειτουργίες:

   α. η δομή να μπορεί να κατασκευαστεί σε χρόνο O(n),
   β. εισαγωγή ή διαγραφή στοιχείου να μπορεί να πραγματοποιηθεί σε χρόνο O(log n),
   γ. είτε το ελάχιστο στοιχείο, ή το μέγιστο στοιχείο να μπορεί να βρεθεί σε σταθερό χρόνο (δηλαδή σε χρόνο O(1)). [15%]

   Σημείωση: Τοποθετήστε τα στοιχεία μέσα στο σορό ως εξής: κάθε κόμβος σε ζυγό βάθος έχει κλειδί μεγαλύτερο από ότι οποιοσδήποτε από τους επιγόνους του, ενώ κάθε κόμβος σε μονό βάθος έχει κλειδί μικρότερο από τους επιγόνους του.

8. Ξεκινώντας με 10 σύνολα S_i, $1 \leq i \leq 10$, καθένα εκ των οποίων περιέχει μόνο ένα στοιχείο, το i, παρουσιάστε τα up-δένδρα που προκύπτουν μετά από την εκτέλεση κάθε μιας από τις ακόλουθες λειτουργίες, δεδομένου ότι χρησιμοποιείται η στρατηγική για μείωση ύψους: A = Union(S_1, S_2); B = Union(A,S_3); C = Union(S_4, B); D = Union(S_5, S_6); E = Union(D, S_7); F = Union(C,E); Find(7);
   Ποια θα ήταν τα up-δένδρα αν η στρατηγική συμπίεσης μονοπατιού είχε επίσης χρησιμοποιηθεί; [10%]

9. α. Αποδείξτε ότι αν δύο κόμβοι ενός γράφου συνδέονται με ένα μονοπάτι, τότε υπάρχει ένα απλό μονοπάτι μεταξύ των δύο αυτών κόμβων.

β. Για κάθε μια από τις 5 ιδιότητες του πρώτου θεωρήματος που παρουσιάζεται στη διαφάνεια 16 της Ενότητας 9, βρείτε ένα γράφο που έχει την αντίστοιχη ιδιότητα αλλά δεν είναι δένδρο.

γ. Εξηγήστε γιατί είναι απαραίτητο ο G να μην είναι πλήρης στην ιδιότητα 3 του δεύτερου θεωρήματος της διαφάνειας 16 της Ενότητας 9. [10%]

10. α. Σε ένα σύνολο αρχείων τα 26 γράμματα του αγγλικού αλφάβητου συναντώνται με την εξής συχνότητα: e 933, t 675, a 571, o 570, i 566, n 537, s 524, r 483, h 349, l 295,c 268, d 255, u 211, m 192, p 183, f 170, g 144, b 120, y 108, w 96, v 71, k 37, x 24, q 11, j 10, z 9. Κατασκευάστε ένα δένδρο Huffman για αυτούς τους 26 χαρακτήρες. Ποια είναι τότε η κωδικοποίηση Huffman της αγγλικής έκδοσης του ονόματός σας;

β. Παρουσιάστε τη λειτουργία του αλγορίθμου συμπίεσης Lempel-Ziv στο ακόλουθο κείμενο (υποθέστε ότι οι γραμμές χωρίζονται με ένα απλό κενό (space)):
abracadabra, abracadab,
abracada, abracad,
abraca, abrac, abra, ab! [10%]

Προγραμματιστικό Μέρος

Να κατασκευάσετε ένα πρόγραμμα που να υλοποιεί τις διαδικασίες αναζήτησης, εισαγωγής και διαγραφής σε ένα πίνακα κατακερματισμού που είναι οργανωμένος με τη μέθοδο:

α. των αλυσίδων

β. της ανοικτής διεύθυνσης με γραμμική αναζήτηση

γ. του ταξινομημένου διπλού κατακερματισμού.

Το κυρίως πρόγραμμα (main()) θα πρέπει να είναι φιλικό προς το χρήστη και με ερωτήσεις να επιτρέπει την επαναληπτική εκτέλεση των αντίστοιχων λειτουργιών. Θα πρέπει να ορίσετε με πρωτοβουλία σας το μέγεθος του πίνακα αν γνωρίζετε ότι ο χώρος κλειδιών περιέχει το πολύ 100 κλειδιά και να είστε σε θέση να δικαιολογήσετε την επιλογή σας. Το κυρίως πρόγραμμα θα πρέπει επίσης να παρέχει τη δυνατότητα αυτόματης εισαγωγής ψευδοτυχαίων αριθμών εκτελώντας κάποια από τις παραπάνω μεθόδους (όποια επιλέξει ο χρήστης) μέχρι ο παράγοντας φόρτου (load factor) να γίνεται ίσος με 0.8. Μετά την εισαγωγή αυτών των αριθμών, το πρόγραμμα θα πρέπει να τυπώνει το ποσοστό των κλειδιών που βρίσκονται στην αρχική τους θέση, τον μέγιστο αριθμό συγκρούσεων και το μέσο αριθμό συγκρούσεων. Ως πρωτεύουσα συνάρτηση κατακερματισμού χρησιμοποιείστε την συνάρτηση διαίρεσης, ενώ ως δευτερεύουσα συνάρτηση κατακερματισμού χρησιμοποιείστε τη συνάρτηση

$$h(K) = 100K^3 + 11K^2 + 79K + 123 \ \% \ m.$$

Το πρόγραμμά σας θα πρέπει να παρέχει συνάρτηση τύπωσης (PrintHashTable()) των περιεχομένων του πίνακα κατακερματισμού, την οποία θα πρέπει ο χρήστης να μπορεί να καλεί από το κυρίως μενού. [100%]

```
ERROR: syntaxerror
OFFENDING COMMAND: --%ztokenexec_continue--

STACK:

false
```