

טבלאות גיבוב – תרגילים עם פתרונות

שאלה 1

- נתונים טבלת גיבוב $T[1..m]$, סדרה של n מפתחות מספריים $\langle k_1, k_2, \dots, k_n \rangle$, ופונקצית גיבוב כלשהי h . נתבונן בתהליך הבא:
- מכניסים את n המפתחות לטבלת הגיבוב על-ידי שימוש בפונקציה h (עם שרשור במקרי התנגשויות);
 - ממיינים כל אחת מהרשימות המקושרות (באמצעות מיון-ערמה);
 - משרשרים את הרשימות המקושרות (לפי סדר התאים בטבלה T).
- א. מהי תוחלת זמן הריצה של תהליך זה?
- ב. מהו זמן הריצה של התהליך במקרה הגרוע?
- ג. האם בכל מקרה מתקבל מיון של סדרת המפתחות $\langle k_1, k_2, \dots, k_n \rangle$? אם לא, באיזה מקרה מתקבלת סדרה ממוינת?

פתרון:

- א. $O(n)$
- ב. $O(n \lg n)$
- ג. הרשימה ממוינת אם $k_1 < k_2 \Rightarrow h(k_1) \leq h(k_2)$

שאלה 2

- בסופרמרקט נמכרים n מוצרים שונים. על כל מוצר מוטבע ברקוד, שהוא מחרוזת בת 10 ספרות. נסמן את קבוצת כל המוצרים ב- S .
- תכננו מבנה נתונים, שיאפשר לבצע על קבוצת המוצרים S את הפעולות הבאות:
- INSERT (S, x): הוספת המוצר x ל- S .
 - DELETE (S, x): הוצאת המוצר x מ- S .
 - PRICE (S, b): החזרת מחירו של המוצר הנושא את הברקוד b .
 - DISCOUNT (S, p): הקטנה ב- $p\%$ של מחירים של כל המוצרים במבנה.

זמן הריצה הממוצע (לכל פעולה): $O(1)$

תארו את מבנה הנתונים המוצע והסבירו בקצרה איך תתבצע כל פעולה.

פתרון:

מבנה הנתונים יורכב מטבלת גיבוב ומשדה D אשר יכיל את ההנחה המצטברת על המוצרים.
פעולת DISCOUNT תתבצע על-יד הכפלת D בהנחה החדשה: $D \leftarrow D * (1 - p/100)$.
פעולת מחיקה תתבצע על-ידי מחיקה מהטבלת גיבוב.
פעולת החזרת מחיר תמצא את המוצר בטבלת גיבוב ותחזיר את המחיר מופל ב-D.
פעולת הכנסה תכניס את המוצר לטבלת הגיבוב כאשר מחירו יחולק ב-D כדי לא להתחשב בירידות המחיר עד כה.

שאלה 3

נתונה טבלת גיבוב $T[1..m]$; כל תא של הטבלה מצביע אל מערך בגודל m . ברצוננו ליישם בכל אחד מ- m המערכים ערימה בינרית. ברצוננו להכניס לטבלת הגיבוב סדרה של m מפתחות בשתי שיטות.
א. בשיטה הראשונה מכניסים את כל m המפתחות למבנה הנ"ל; אחרי שכולם הוכנסו, בונים את m הערימות. מהו זמן הריצה של השיטה במקרה הגרוע? האם יתכן שמתקבל זמן ריצה טוב יותר בממוצע?
ב. בשיטה השנייה מכניסים את m המפתחות למבנה הנ"ל; אחרי כל הכנסת מפתח, מתקנים את הערימה המתאימה. מהו זמן הריצה של השיטה במקרה הגרוע ובממוצע?

פתרון:

נניח ש $n = \Theta(m)$, כנדרש מטבלת גיבוב.

א. המקרה הגרוע $O(n)$: כל המפתחות נכנסים לאותו תא. בממוצע בכל תא יהיו $O(1)$ מפתחות אך זמן הריצה יהיה עדיין $O(n)$.
ב. במקרה הגרוע $O(n \lg n)$: כל המפתחות נכנסים לאותו תא. בממוצע בכל תא יהיו $O(1)$ מפתחות ולכן זמן הריצה יהיה $O(n)$.

שאלה 4

פתרו את תרגיל 12.1-4 מהספר של קורמן.

פתרון:

נתאר מימוש של מערך A שניתן לבצע עליו אתחול, קריאה וכתיבה בזמן קבוע.

המערך A מורכב מחמישה שדות:

A.value – הוא מערך באורך n המכיל את האיברים המאוחסנים ב-A

A.stack – הוא מערך באורך n המשמש כמחסנית לאחסון אינדקסי התאים שערכם אינו הערך

המאוחסן בהם (A.value[i]), אלא ערך האתחול (A.init)

A.pointer – הוא מערך באורך n המכיל את האינדקס של התא במחסנית המכיל את האינדקס של

התא במערך זה (אם קיים תא כזה). זה קצת מבלבל אבל יובהר בהמשך.

A.top – האינדקס של ראש המחסנית A.stack.

A.init – ערך האתחול.

כיצד מתבצע אתחול המערך לערך x?

המערך מאותחל על-ידי הצבת הערך x ב-A.init ואיפוס A.top.

INIT(A,x)

A.init ← x

A.top ← 0

כיצד מתבצעת קריאת ערך מאינדקס i במערך?

קודם כל האלגוריתם יבדוק אם הערך הכתוב הוא הערך "האמיתי" או שצריך להשתמש בערך

האתחול: הערך הכתוב הוא הערך האמיתי אם i האינדקס i נמצא במחסנית. לאחר הבדיקה ידע

האלגוריתם איזה ערך להחזיר: ערך האתחול או הערך שבמערך.

READ(A,i)

if A.top > 0 and A.pointer[i] ≤ A.top and A.stack[A.pointer[i]] = i

then return A.value[i]

else return A.init

כיצד מתבצעת כתיבת ערך לאינדקס i?

קודם כל יבדוק האלגוריתם אם כבר בוצעה כתיבה לתא i. אם כן הערך יוצב בתא i וזהו. אם וז

הכתיבה הראשונה לתא, מאז האתחול האחרון, על האלגוריתם לבצע את הפעולות הבאות: להכניס

את הערך לתא, לדחוף את האינדקס למחסנית (מה שיסמן שערך נכתב לתא) ולעדכן מצביע לתא

המתאים במחסנית.

WRITE(A,i,x)

A.value[i] ← x

```
if not (A.top>0 and A.pointer[i]≤A.top and  
    A.stack[A.pointer[i]]=i)  
then A.top←A.top+1  
    A.stack[A.top]←i  
    A.pointer[i]←A.top
```

שאלה 5

נתונות שתי רשימות של מספרים, אחת בת m איברים והשניה בת n איברים. תארו אלגוריתם הקובע האם קיים איבר משותף לשתי הרשימות. זמן הריצה של האלגוריתם $O(\max(m,n))$ בממוצע.

פתרון:

נכניס את האיברים מהרשימה הראשונה לטבלת גיבוב. לכל איבר מהרשימה הראשונה נבדוק האם הוא מופיע בטבלת הגיבוב. זמן הריצה, בממוצע, הוא $O(m+n)$. אבל $m+n \leq 2\max(n,m)$ ולכן זמן הריצה, בממוצע, הוא $O(\max(m,n))$, כנדרש.

שאלה 6

נתונה קבוצת איברים S . על הקבוצה S מוגדרות הפעולות הבאות:
INSERT(S, x) - הכנסת האיבר x ל- S .
SEARCH(S, k) - החזרת מצביע לאיבר ב- S בעל מפתח k ;
אם אין ב- S איבר כזה - החזרת NIL.

הקבוצה S מוגבלת בגודלה, ולכן לפני כל ביצוע פעולה של הכנסת איבר ל- S צריך למחוק איבר מ- S כדי לפנות מקום לאיבר חדש.

מחיקת האיבר יכולה להתבצע באחת משתי שיטות:

FIFO (First In First Out) - מחיקת האיבר שהוכנס ל-S לפני הכי הרבה זמן.

LRU (Least Recently Used) - מחיקת האיבר ב-S שניגשנו אליו לפני הכי הרבה זמן.

(כל פעולה של החזרת מצביע לאיבר x וכל פעולה של הכנסת x ל-S הן פעולות שבהן התבצעה גישה לאיבר x).

מכיוון שיש חשיבות רבה למהירות הביצוע של הפעולות השונות, כל פעולה צריכה להתבצע בזמן $O(1)$ בממוצע.

הציעו מבנה נתונים עבור הקבוצה S כאשר:

א. מחיקת איבר מתבצעת בשיטת FIFO.

ב. מחיקת איבר מתבצעת בשיטת LRU.

פתרון:

א. טבלת גיבוב + תור. לפני כל הכנסה, אם צריך למחוק איבר יימחק האיבר שבראש התור.
ב. טבלת גיבוב + רשימה מקושרת דו-כיוונית. מכל איבר בטבלה יהיה מצביע לאיבר ברשימה המקושרת. כל איבר שניגשים אליו יועבר לתחילת הרשימה. בהכנסה, אם צריך למחוק איבר, יימחק האיבר בסוף הרשימה.

שאלה 7

הציעו מבנה נתונים S התומך בפעולות הבאות:

SEARCH (S, k): חיפוש אחר המפתח k במבנה S, והחזרת איבר כלשהו בעל מפתח k ;

INSERT (S, k): הכנסת איבר חדש בעל המפתח k למבנה S;

DELETE (S, k): מחיקת איבר כלשהו בעל המפתח k מהמבנה S;

FREQUENCY (S, k): החזרת מספר האיברים בעלי המפתח k שבמבנה S;

זמן הריצה הנדרש של כל אחת מהפעולות הינו $O(1)$ בממוצע.

פתרון:

מבנה הנתונים הוא טבלת גיבוב שלכל צמת המייצג מפתח מסוים יש מצביע לרשימה מקושרת לכל האיברים בעלי אותו המפתח. בצמת יהיה, בנוסף, מונה לאורך הרשימה המקושרת.

הכנסה, ייבדק האם קיים איבר בעל מפתח k בטבלה. אם כן – יתווסף האיבר לרשימה המקושרת המתאימה ומונה מספר איברי הרשימה יגדל ב-1. אם לא – תתווסף רשימה מקושרת חדשה שגודלה 1.

במחיקה, אם יש איבר בעל מפתח k , יימחק האיבר הראשון ברשימה המתאימה, וגודלה יקטן ב-1. אם גודל הרשימה התאפס, יימחק המפתח מן הטבלה. בחיפוש יוחזר האיבר הראשון ברשימה המקושרת של המפתח k (אם יש רשימה כזאת) כדי לדעת את מספר האיברים בעלי המפתח k , ייקרא המונה של הרשימה של המפתח k . אם אין רשימה כזאת מספר המופעים הוא 0.

שאלה 8

Consider the following algorithm on a list of n integers:

1. Insert the integers into a hash table
2. Sort the linked list in every slot
3. Concatenate the linked lists in the slots (in ordered fashion, from slot 0 to slot $m-1$) into one long linked list.

- a. What is the expected running time of this algorithm? Explain.
- b. What condition should the hash function satisfy, if we want the linked list created, in step 3, to be a sorted list? Prove your answer

פתרון:

א. צעד 1 מתבצע בזמן קבוע לכל איבר. מכיוון שבממוצע יש $O(1)$ ערכים בכל תא, אז זמן מיון כל רשימה הוא קבוע (לא משנה באיזה אלגוריתם נשתמש). סה"כ זמן הריצה $O(m+n)$ כש- m הוא מספר התאים בטבלה ו- n הוא מספר האיברים.

ב. אם $k_1 < k_2$ אז $h(k_1) \leq h(k_2)$