

תרגיל אחרון במבנה נתונים

שאלה 1

נתון מערך A המכיל n מספרים שלמים. בהינתן מספר x המופיע ב A , נגדיר את הריבוי של x במערך A בתור מספר המופיעים של x ב- A . רוצים לייצר מערך ממונן B בגודל n , שיכיל את הריבוי של כל אחד מהמספרים המופיעים ב A . הציעו אלגוריתם שפותר את הבעיה **בתוחלת** זמן הטובה ביותר.

שאלה 2

נתונה טבלת גיבוב בגודל m , $m > 3$, ראשוני. נשתמש בשיטת הבדיקה הריבועית בעזרת פונקציית הגיבוב $h(k, i) = (\text{hash}(k) + i^2) \bmod m$, $i \geq 0$.
א' הוכיחו שאם הטבלה פחות מחצי מלאה, אזי תמיד נמצא מקום עבור מפתח חדש.
ב' הראו, באמצעות דוגמה נגדית, שאם הטבלה יותר מחצי מלאה, אזי קיימת האפשרות שלא נוכל למצוא מקום עבור מפתח חדש.

שאלה 3

1. תהי H טבלת גיבוב עם chaining ועם universal hashing. נניח ש- H היא בעלת $n^{3/2}$ תאים, ושמתכניסים אליה קבוצת איברים S בעלת n איברים **שוניים**. נסמן ב- h את פונקציית ה-hash.
א. מה תוחלת מספר ההתנגשויות? (תזכורת: "ההתנגשות" היא זוג איברים x, y כך ש- $h(x)=h(y)$).
ב. הוכיחו שההסתברות שמס' ההתנגשויות גדול מ- $2\sqrt{n}$ היא $\Omega(1)$.

שאלה 4

תארו אלגוריתם לפתרון הבעיה הבאה:
קלט: קבוצה כלשהי $A = \{a_1, \dots, a_n\}$ של n מספרים.
מטרה: האלגוריתם צריך לבנות **בזמן ממוצע טוב ביותר** מבנה נתונים DS בעל התכונה הבאה:
בהינתן אוסף S של m מספרים שאיבריו לקוחים מ A , ניתן להיעזר ב DS בכדי למיין את איברי S בזמן $O(\max(m, n))$ ב **WC**. שימו לב כי S יכולה להכיל איבר כלשהו של A **מספר פעמים**. על כן ייתכן כי מתקיים $m > n$.

שאלה 5

במחשב אישי קיימים בד"כ זיכרון חיצוני (דיסק קשיח) וזיכרון פנימי. הגישה לזיכרון הפנימי מהירה יותר אך גודלו קטן יותר מהזיכרון החיצוני. לכן, נתונים שנקראים לעיתים קרובות כדאי להעתיק מהדיסק הקשיח אל הזיכרון הפנימי כדי לקצר את זמני הגישה. לצורך בניית מנגנון כזה (המכונה Cache) נגדיר את מבנה הנתונים הבא. נתונה קבוצה S של איברים. על הקבוצה S מוגדרות הפעולות הבאות:

- INSERT (S, x) - הכנסת האיבר x ל- S .
- SEARCH (S, x) - החזר מצביע לאיבר ב- S בעל מפתח k ;
אם אין ב- S איבר כזה - החזר nil.

הקבוצה S היא מוגבלת בגודלה, ולכן לפני ביצוע פעולה של הכנסת איבר ל- S צריך למחוק איבר מ- S בכדי לפנות מקום לאיבר החדש. מחיקת האיבר יכולה להתבצע באחת משתי שיטות:

1. FIFO (First In First Out) - מחיקת האיבר שהוכנס ל- S לפני הכי הרבה זמן.
2. LRU (Least Recently Used) - מחיקת האיבר ב- S שניגשנו אליו לפני הכי הרבה זמן (כל פעולה של החזרת מצביע לאיבר x , או הפעולה של הכנסת x ל- S הן פעולות שבהן התבצעה גישה ל- x).

מכיוון שיש חשיבות רבה למהירות הביצוע של הפעולות השונות, כל פעולה צריכה להתבצע בזמן $O(1)$ בתוחלת. הצע מבנה נתונים עבור הקבוצה S כאשר:

- א. מחיקת איבר מתבצעת בשיטת FIFO.
- ב. מחיקת איבר מתבצעת בשיטת LRU.