

אוניברסיטת ת"א סמסטר א' תש"ע

מבני נתונים – תרגיל מס. 8

1. בכיתה ראינו אלגוריתם ליניארי המממש פעולת select ע"י חלוקת הקלט לחמישיות.
א. האם האלגוריתם ירוץ בזמן ליניארי גם אם יחולקו איברי הקלט לקבוצות בנות 7 איברים?
ב. מדוע זמן הריצה של SELECT אינו ליניארי כאשר החלוקה היא לקבוצות בנות 3 איברים?
2. נתון מערך בגודל n ללא ערכים כפולים ושני מספרים שלמים k_1 ו- k_2 המקיימים $1 \leq k_1 \leq k_2 \leq n$.
א. תארו אלגוריתם שמחזיר את כל האיברים שדרגתם במערך בין k_1 ל- k_2 (כלומר האיברים שדרגתם (k_1, k_1+1, \dots, k_2) בזמן ריצה $O(n)$ (דרגה של איבר במערך היא מספר האיברים הקטנים ממנו או שווים לו).
ב. נגדיר $k = k_2 - k_1$. מה זמן הריצה האופטימלי כפונקציה של n ו- k שניתן להשיג אם דורשים שהאיברים יוחזרו ממוינים?
3. תהי H טבלת Hash בגודל n שמשמשת ב-rehash ואינה מאפשרת מחיקות. נוסף לכל תא i בטבלה מונה c_i שסופר את המפתחות k שהוכנסו לטבלה ומקיימים $h_1(k) = i$.
א. כיצד ניתן להשתמש במונים כדי להקטין את מספר הגישות לתאים בטבלה בחיפוש כושל?
ב. תנו דוגמה לחיפוש כושל שעבורו מספר הגישות לתאים בטבלה קטן מ- n ל-2 (ע"י שימוש בהצעה מסעיף א').
4. משפחת פונקציות H מ- U ל- $\{0, 1, \dots, m-1\}$ היא 100-weakly universal אם לכל x ו- y ב- U השונים זה מזה מתקיים
$$\Pr_{h \in H}[h(x) = h(y)] \leq 100/m.$$

תהי $U = \{0, 1, \dots, u-1\}$ קבוצת מפתחות כך ש- $m > 100$ ו- $u \gg m$.
נגדיר את הפונקציה f_a כך:
$$f_a(x) = ax \bmod m.$$

הראו כי מחלקת הפונקציות $H = \{f_a \mid 0 < a < m\}$ איננה 100-weakly universal.
5. נתונה טבלת Hash בגודל m המשמשת לאחסון n איברים, $n \leq m/2$. התנגשויות נפתרות ע"י open addressing.
א. בהנחת ה-uniform hashing, הראו כי עבור $n, i = 1, 2, \dots$, ההסתברות שמספר הבדיקות הדרושות להכנסה ה- i גדול ממש k היא לכל היותר 2^{-k} .
ב. הראו שעבור $n, i = 1, 2, \dots$, ההסתברות שההכנסה ה- i דורשת יותר מאשר $2 \log n$ בדיקות היא לכל היותר $1/n^2$.
ג. יהי X_i משתנה מקרי המציין את מספר הבדיקות הדרושות להכנסה ה- i . בסעיף ב' הראיתם כי $\Pr[X_i > 2 \log n] \leq 1/n^2$. יהי X מ"מ המציין את מספר הבדיקות המקסימאלי הדרוש להכנסה כלשהי מבין n ההכנסות. הראו כי
$$\Pr[X > 2 \log n] \leq 1/n.$$

השתמשו לשם כך בחסם האיחוד -

$$P\left(\bigcup_i A_i\right) \leq \sum_i P(A_i).$$