מבני נתונים 67109
תרגיל 8

להגשה: עד ה-22/5/2019

שאלה 1 - לשחק קצת עם קוד האפמן

1. עבור המחרוזת הבאה, מעל הא"ב $\{$ש,ר,ה,י,מ,ח$\} = \Sigma$, חשבו את השכיחויות של כל אות בה וייצרו את קוד האפמן שלה - ציירו את העץ וחשבו את תוחלת אורך הקידוד עבורו. כתבו את הקידוד של המחרוזת. כמה ביטים יש בה?
שרהשרהשירשמחשירשמחשרהשרה
(כאשר המחרוזת נקראת מימין לשמאל. הקידוד יקרא משמאל לימין. כלומר אם לדוגמא "ש" יקודד ע"י 01 ו-"ר" ע"י 00 אז הקידוד יתחיל 0100 וימשיך לימין זוג האפסים)

2. כעת קודדו את אותה המחרוזת כאשר הא"ב איתו אתם עובדים הוא של זוגות של אותיות. רשמו מה הא"ב שלכם באופן מפורש (אין צורך לקחת זוגות שאינן מופיעות) וחשבו את השכיחויות של כל אות בו. ציירו את העץ לקוד האפמן של המחרוזת עם הא"ב הזה והשכיחויות שחישבתם וחשבו את תוחלת אורך הקידוד עבורו. כתבו את הקידוד של המחרוזת. כמה ביטים יש בה?

3. בפעם האחרונה - קודדו את אותה המחרוזת כאשר הא"ב איתו אתם עובדים הוא של שלשות של אותיות. (חשבו שכיחויות גם כאן, ציירו עץ עבור הא"ב עם השכיחויות שחישתבם וחשבו את תוחלת אורך הקידוד. כתבו את הקידוד של המחרוזת ואת אורכו בביטים).

שאלה 2 - מקרה מיוחד בקוד האפמן

יהי T עץ האפמן עבור א"ב Σ ופונקציית שכיחויות $p : \Sigma \to (0,1]$. הוכיחו כי אם קיימת אות $\sigma \in \Sigma$ כך ש-$p(\sigma) > 0.4$ אזי קיימת אות $\tau \in \Sigma$ שמקודדת תחת T ע"י ביט בודד (כלומר שהקידוד שלה הוא מחרוזת באורך של 1).

שאלה 3 - לשחק קצת עם פונקציות גיבוב

יהיו $M = 11$, $h_1(x) = x \mod M$, $h_2(x) = \left\lfloor \frac{x}{M} \right\rfloor mod \quad M+1$, $c_1 = 1$, $c_2 = 2$.
הציגו הכנסה של המספרים הבאים (משמאל לימין) 57,76,63,77,54,31,23,74 לטבלאת גיבוב בגודל M כאשר התנגשויות מנוהלות בעזרת 3 השיטות שראינו:

- שיטת Linear probing בשימוש $h(x,i) = h_1(x) + i(\mod M)$

- שיטת Quadratic probing בשימוש $h(x,i) = h_1(x) + c_1 i + c_2 i^2 (\mod M)$
- שיטת Double hashing בשימוש $h(x,i) = h_1(x) + i h_2(x) (\mod M)$

כמה התנגשויות קרו בשימוש בכל אחת מהשיטות? האם זה מקרי?
בתשובתכם הראו רק את התוצאה הסופית של המערך עבור כל שיטה ורשמו כמה התנגשויות קרו בשימושה.

שאלה 4 - שימוש בפונקציות גיבוב

יהי U קבוצה סופית של טבעיים.

1. תארו מבנה נתונים שצורך $O(n)$ מקום בזיכרון ותומך בפעולות הבאות:

- אתחול $Init(S)$: בהנתן תת-קבוצה של מספרים שונים $S \subseteq U$, $|S| = n$, הפעולה מאתחלת את מבנה הנתונים בזמן ריצה של $O(f(n))$ עבור פולינום f כלשהו.
- בדיקת שייכות $Member(x)$: בהנתן איבר $x \in U$, הפעולה מחזירה $True$ אם $x \in S$ ואחרת היא מחזירה $False$ בזמן ריצה ממוצע $O(1)$.
- בחירה $Select(k)$: מחזירה את האיבר ה-k בגודלו (שיש בדיוק $k-1$ קטנים ממנו) מ-S בזמן ריצה $O(1)$.

2. בהנחה שכלל פעולות מבנה הנתונים מבוססות על השוואות, האם ניתן לבנות מבנה נתונים שיתמוך בפעולות הוצאה $Delete(x)$ והכנסה $Insert(x)$, שניהם בזמן ריצה $O(1)$ (בנוסף לפעולות המצויינות לעיל)? אם כן, תארו מבנה נתונים כזה. אחרת, הוכיחו שלא ניתן לבנות כזה.