פתרון תרגיל מספר 4 - דאסט

שם: מיכאל גרינבאום, **ת"ז:** 211747639

6 באפריל 2019

1. **פתרון:**

(א) **צ"ל:** האלגוריתם הוא יציב
הוכחה:
יהי $0 \le l \le k$, ונסמן ב$i_1, \dots, i_m$ את האינדקסים כך שמתקיים $S[i_1] = \cdots = A[i_m] = l$,
נסמן בl_j את $a[i_j] = l = l_j$, וב$j_1, \dots, j_m$ את המיקום של $A[i_j]$ במערך B
ונראה שלאחר הפעלת האלגוריתם, מתקיים $j_1 < \cdots < j_m$, ומכך נסיק כי האלגוריתם יציב,
יהי אינדקס $2 \le q \le m$ ונראה כי $j_{q-1} < j_q$:
נסתכל על הרצת האלגוריתם, אחרי הלופ הראשון מתקיים כי $C[a[i_q]] = C[a[i_{q-1}]] = m$,
בלופ השני מתקיים כי $C[a[i_q]] = C[a[i_{q-1}]] = C[l-1] + m$,
נשים לב כי בלופ השלישי ל$j = i_q$ מתקיים $B[C[A[i_q]]] = B[C[l]] = B[C[l-1] + m - (m-q)]$ כי הוא מגיע ל$C[l]$, $m-q$ פעמים לפני שמגיע לq, ולכן מחסר $m-q$ מהתוצאה ההתחלתית,
לכן לכל q, מתקיים כי $j_q = C[l-1] + m - (m-q)$, לכן

$$j_{q-1} < j_q \Leftrightarrow C[l-1] + m - (m-q-1) < C[l-1] + m - (m-q) \Leftrightarrow -1 < 0$$

כלומר קיבלנו כי $j_{q-1} < j_q$ תמיד מתקיים ולכן $j_1 < \cdots < j_m$, כלומר האלגוריתם שומר על סדר ההופעה.
כלומר האלגוריתם יציב

מ.ש.ל.א.☺

(ב) **צ"ל:** האלגוריתם הוא אינו יציב ונכון
הוכחה:
תחילה נראה שהוא לא יציב, נגדיר $A = [a_0 = 1, a_1 = 1]$, נשים לב כי לאחר הפעלת האלגוריתם נקבל כי $B = [a_1 = 1, a_0 = 1]$, כלומר האלגוריתם לא שומר על סדר ההופעה, לכן האלגוריתם לא יציב.
עתה נראה מדוע האלגוריתם נכון, נשים לב שהשוני היחיד בין האלגוריתם בסעיף זה ביחס לסעיף הקודם הוא סדר הריצה על המערך בלופ האחרון (וסדר הריצה בשני הלופים הראשונים לא משנה כלום),
לכן $\text{counting\_sort2}(A[1:n]) = \text{counting\_sort1}(A[n:1])$, לכן מנכונות האלגוריתם בסעיף הקודם, נקבל כי האלגוריתם בסעיף זה גם הוא עובד ומחזיר מערך ממוין

מ.ש.ל.ב.☺

(ג) **צ"ל:** האלגוריתם הבא נכון וביעילות $\Theta(n+k)$
הוכחה:

```
def countAtoB(A, k, a, b):
    C = [0] * (k+1)
    for j in range(0,len(A)):
        C[A[j]] += 1
    for i in range(1, len(C)):
        C[i] += C[i-1]
    if a == 0:
        return C[b]
    return C[b] - C[a-1]
```

נוכיח בשלבים,

i. **שלב ראשון:** לאחר הריצה על הלולאה הראשונה $C[i]$ מכיל את מספר ההופעות של i בA,
נוכיח זאת בשמורת לולאה,
<u>בסיס:</u> נשים לב כשאנחנו נכנסים ללולאה, המערך מכיל 0 הופעות לכל i, כמו שהיינו מצפים במקרה הריק.

צעד לולאה: נניח כי C מכיל את מספר ההופעות של i (לכל i) במערך $A\left[1:j\right]$ ונראה שלאחר הרצה של הלולאה C מכיל את מספר ההופעות של i (לכל i) במערך $A\left[1:j+1\right]$,
נשים לב כי מספר ההופעות של $A\left[j+1\right]$ במערך $A\left[1:j+1\right]$ שווה ל1 + מספר ההופעות במערך $A\left[1:j\right]$ ומספר זה שווה ל$C\left[j+1\right]+1$, ואותו אנחנו מכניסים ל$C\left[j+1\right]$, כנדרש,
ולשאר המספרים ב$A\left[1:j+1\right]$ מספר ההופעות שלהם שווה למספר ההופעות שלהם ב$A\left[1:j\right]$ (כי אינם מופיעים ב$A\left[j+1\right]$) ולכן $C\left[i\right]$, לא מתעדכן ולכן מכיל עדיין את מספר ההופעות של i ב$A\left[1:j\right]$ שווה למספר ההופעות של i ב$A\left[1:j+1\right]$

ii. **שלב שני:** לאחר הריצה על הלולאה השנייה $C\left[i\right]$ מכיל את מספר ההופעות של כל המספרים קטנים שווים לi,
נוכיח זאת בשמורת לולאה,
בסיס: נשים לב כשאנחנו נכנסים ללולאה, המערך מכיל 0 הופעות לכל i, התנאי נכון באורך ריק כי אין מספרים קטנים מ0 מההנחה
צעד לולאה: נניח כי $C\left[i\right]$ מכיל את מספר ההופעות של המספרים קטנים שווים לi מערך $A\left[1:n\right]$ ונראה שלאחר הרצה של $C\left[i+1\right]$ מכיל את מספר ההופעות של המספרים קטנים שווים ל$i+1$,
נשים לב כי $C\left[i+1\right]$ מכיל את מספר האיברים השווים ל$i+1$ ואילו $C\left[i\right]$ מכיל את מספר האיברים הקטנים מ$i+1$, לכן $C\left[i+1\right]=C\left[i\right]+C\left[i+1\right]$ מכיל את מספר ההופעות של המספרים קטנים שווים ל$i+1$, כנדרש

לכן בסוף הריצה $C\left[i\right]$ מכיל את מספר ההופעות של כל המספרים קטנים שווים לi,
אם $a=0$, נרצה לדעת את כמות המספרים קטנים שווים b, לכן נחזיר $C\left[b\right]$ (כמו שהאלגוריתם עושה)
אחרת נרצה את כמות המספרים בין a לb - והוא שווה לכמות המספרים הקטנים שווים b **פחות** כמות המספרים שקטנים ממש מa, לכן האלגוריתם אמור להחזיר $C\left[b\right]-C\left[a-1\right]$ (כמו שהאלגוריתם עושה)

עתה, נשים לב כי $n+k\leq T\left(n\right)=k+n+\left(k-1\right)+3=n+2k+2\leq3\left(n+k\right)$,
לכן $T\left(n\right)=\Theta\left(n+k\right)$

מ.ש.ל.ג.☺

2. **פתרון:**

(א) **צ"ל:** $R\left(n\right)\leq U\left(n\right)$
הוכחה:
נוכיח באינדוקציה על n, תחילה נשים לב כי $\exists c'>0$ כך ש$\forall n\in\mathbb{N}$ מתקיים $R\left(n\right)\leq\frac{2}{n}\sum_{m=1}^{n-1}R\left(m\right)+c'\cdot n$,
נגדיר $c=\max\left\{c',R\left(1\right)\right\}$
בסיס: $n=1$,

$$R\left(1\right)\leq\max\left\{c',R\left(1\right)\right\}\cdot1=c\cdot1=U\left(1\right)$$

צעד: נניח שהטענה נכונה ל$1,\ldots,n-1$ ונראה שהיא נכונה לn

$$R\left(n\right)\leq\frac{2}{n}\sum_{m=1}^{n-1}R\left(m\right)+c'\cdot n\leq R\left(n\right)\leq\frac{2}{n}\sum_{m=1}^{n-1}R\left(m\right)+c\cdot n\overset{\text{induction}}{\leq}\frac{2}{n}\sum_{m=1}^{n-1}U\left(m\right)+c\cdot n=U\left(n\right)$$

כלומר הטענה נכונה לn, כנדרש

מ.ש.ל.א.☺

(ב) **צ"ל:** $\lim_{n\to\infty}\left[\sum_{i=1}^{n}\frac{\left(-1\right)^{i+1}}{i}\right]=\ln\left(2\right)$
הוכחה:

נשים לב כי

$$\begin{aligned}\lim_{n\to\infty}\left[\sum_{i=1}^{n}\frac{(-1)^{i+1}}{i}\right] &= \lim_{n\to\infty}\left[\sum_{i=1}^{2n}\frac{1}{i}-2\cdot\sum_{i=1}^{n}\frac{1}{2i}\right]=\lim_{n\to\infty}\left[\sum_{i=1}^{2n}\frac{1}{i}+\ln(2n)-\ln(2n)-2\cdot\sum_{i=1}^{n}\frac{1}{2i}\right]\\ &\overset{\star}{=}\gamma+\lim_{n\to\infty}\left[\ln(2n)-2\cdot\sum_{i=1}^{n}\frac{1}{2i}\right]=\gamma+\lim_{n\to\infty}\left[\ln(n)+\ln(2)-2\cdot\sum_{i=1}^{n}\frac{1}{2i}\right]\\ &\overset{\star}{=}\gamma+\ln(2)+\lim_{n\to\infty}\left[\ln(n)-2\cdot\sum_{i=1}^{n}\frac{1}{2i}\right]=\gamma+\ln(2)+\lim_{n\to\infty}\left[\ln(n)-\sum_{i=1}^{n}\frac{1}{i}\right]\\ &\gamma+\ln(2)-\gamma=\ln(2)\end{aligned}$$

ב$\star$ הנחנו שהגבולות קיימים ולאחר מכן הראנו שהם קיימים, לכן אין בעיה עם המעברים שנעשו וקיבלנו כי

$$\boxed{\lim_{n\to\infty}\left[\sum_{i=1}^{n}\frac{(-1)^{i+1}}{i}\right]=\ln(2)}$$

מ.ש.ל.ב.☺

3. **פתרון:**

(א) **צ"ל:** $\mathbb{E}(X)$

הוכחה:

נגדיר $X(\omega=k)=\max\{n-k,k\}$, נגדיר $X_i(\omega=k)=\begin{cases}\max\{n-k,k\} & i=k\\ 0 & i\neq k\end{cases}$,

נשים לב כי $X=\sum_{i=0}^{n}X_i$, לכן

$$\begin{aligned}\mathbb{E}(X) &= E\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=0}^{n}\mathbb{E}(X_i)=\sum_{i=0}^{n}\left(\sum_{\omega\in\Omega}X_i(\omega)\cdot\mathbb{P}(\omega)\right)\\ &=\sum_{i=0}^{n}\left(\sum_{\omega\in\Omega,\omega=i}X_i(\omega)\cdot\mathbb{P}(\omega)\right)=\sum_{i=0}^{n}\left(\max\{n-i,i\}\cdot\frac{1}{n}\right)\\ &=\frac{1}{n}\cdot\sum_{i=0}^{n}\max\{n-i,i\}=\frac{1}{n}\cdot\left(\sum_{i=0}^{\lfloor\frac{n}{2}\rfloor}\max\{n-i,i\}+\sum_{i=\lfloor\frac{n}{2}\rfloor+1}^{n}\max\{n-i,i\}\right)\\ &=\frac{1}{n}\cdot\left(\sum_{i=0}^{\lfloor\frac{n}{2}\rfloor}(n-i)+\sum_{i=\lfloor\frac{n}{2}\rfloor+1}^{n}i\right)=\frac{2}{n}\cdot\left[\sum_{i=0}^{n}i-\sum_{i=0}^{\lfloor\frac{n}{2}\rfloor}i\right]=\frac{2}{n}\cdot\left(\frac{n(n+1)}{2}-\frac{\lfloor\frac{n}{2}\rfloor\left(\lfloor\frac{n}{2}\rfloor+1\right)}{2}\right)\\ &=(n+1)-\frac{\lfloor\frac{n}{2}\rfloor\left(\lfloor\frac{n}{2}\rfloor+1\right)}{n}=\begin{cases}2l+1-\frac{l+1}{2} & n=2l\\ 2l-\frac{l\cdot(l-1)}{2l-1} & n=2l-1\end{cases}=\begin{cases}\frac{3}{2}l+\frac{1}{2} & n=2l\\ l\cdot\left[2-\frac{(l-1)}{2l-1}\right] & n=2l-1\end{cases}\\ &=\begin{cases}\frac{3}{4}n+\frac{1}{2} & n=2l\\ l\cdot\left[\frac{(4l-2)-(l-1)}{2l-1}\right] & n=2l-1\end{cases}=\begin{cases}\frac{3}{4}n+\frac{1}{2} & n=2l\\ l\cdot\frac{3l-1}{2l-1} & n=2l-1\end{cases}\\ &=\begin{cases}\frac{3}{4}n+\frac{1}{2} & n=2l\\ \frac{n+1}{2}\cdot\frac{3\left(\frac{n+1}{2}\right)-1}{n} & n=2l-1\end{cases}=\begin{cases}\frac{3}{4}n+\frac{1}{2} & n=2l\\ \frac{n+1}{2}\cdot\frac{\frac{3n}{2}-\frac{1}{2}}{n} & n=2l-1\end{cases}\\ &=\begin{cases}\frac{3}{4}n+\frac{1}{2} & n=2l\\ \frac{n+1}{4}\cdot\frac{3n-1}{n} & n=2l-1\end{cases}\end{aligned}$$

לכן

$$\boxed{\mathbb{E}\left(X\right)=\begin{cases}\frac{3}{4}n+\frac{1}{2} & n=2l\\ \frac{n+1}{4}\cdot\frac{3n-1}{n} & n=2l-1\end{cases}\approx\frac{3}{4}n}$$

מ.ש.ל.א.☺

(ב) **צ"ל:** נמקו מדוע $\frac{3}{4}n$ הוא הגודל הממוצע ולמה אי אפשר להשתמש בו בנוסחא הרקורסיבית
הוכחה:
מסעיף א' ראינו ש$\frac{3}{4}n$ הוא הערך הממוצע לחצות מערך ל2 חלקים כך שאחד הוא הגדול מהשני.
במקרה הגרוע אנחנו תמיד נקבל את החלק הגדול והוא יהיה $\frac{3}{4}n$ בממוצע.
הסיבה שאי אפשר להשתמש ב$\frac{3}{4}n$ בנוסחא הרקורסיבית היא שיש מצבים יותר גרועים, אך בממוצע זה ייצא $\frac{3}{4}n$

מ.ש.ל.ב.☺

(ג) **צ"ל:** $\mathbb{E}\left(f\left(i\right)\right)>\max\left\{ f\left(\left\lfloor \mathbb{E}\left[i\right]\right\rfloor \right),f\left(\left\lceil \mathbb{E}\left[i\right]\right\rceil \right)\right\}$
הוכחה:
נגדיר $f\left(\omega\right)=\begin{cases}2 & \omega=2\\ 3 & \omega=1,3\end{cases}$, עבור $n=3$ והסתברות אחידה כלומר $\mathbb{P}\left(\omega\right)=\frac{1}{n}$, לכן

$$f\left(\left\lceil \mathbb{E}\left[i\right]\right\rceil \right)=f\left(\left\lceil \sum_{\omega\in\Omega}\omega\cdot\mathbb{P}\left(\omega\right)\right\rceil \right)=f\left(\left\lceil \frac{1}{n}\cdot\sum_{i=1}^{n}i\right\rceil \right)=f\left(\left\lceil \frac{n+1}{2}\right\rceil \right)=f\left(2\right)=2$$

$$f\left(\left\lfloor \mathbb{E}\left[i\right]\right\rfloor \right)=f\left(\left\lfloor \sum_{\omega\in\Omega}\omega\cdot\mathbb{P}\left(\omega\right)\right\rfloor \right)=f\left(\left\lfloor \frac{1}{n}\cdot\sum_{i=1}^{n}i\right\rfloor \right)=f\left(\left\lfloor \frac{n+1}{2}\right\rfloor \right)=f\left(2\right)=2$$

$$\mathbb{E}\left(f\left(i\right)\right)=\sum_{\omega\in\Omega}f\left(\omega\right)\cdot\mathbb{P}\left(\omega\right)=\frac{1}{3}\cdot\left(\sum_{\omega\in\Omega}f\left(\omega\right)\right)=\frac{1}{3}\cdot\left(2+3+3\right)=2+\frac{2}{3}$$

כלומר

$$\boxed{\mathbb{E}\left(f\left(i\right)\right)=2+\frac{2}{3}>2=\max\left\{ f\left(\left\lfloor \mathbb{E}\left[i\right]\right\rfloor \right),f\left(\left\lceil \mathbb{E}\left[i\right]\right\rceil \right)\right\}}$$

מ.ש.ל.ג.☺

4. **פתרון:**

(א) **צ"ל:** יש לפחות $\frac{3}{10}n$ קטנים מp ולפחות $\frac{3}{10}n$ גדולים מp
הוכחה:
מהיות p הוא האמצעי של האמצעים, יש לפחות $\frac{n}{10}$ אמצעים קטנים או שווים ממנו (כי הסתכלנו על $\frac{n}{5}$ אמצעים), ועבור כל אחד מהאמצעים הללו יש **בדיוק** 3 איברים קטנים או שווים ממנו, לכן יש לפחות $3\cdot\frac{n}{10}$ איברים קטנים או שווים לp.
מהיות p הוא האמצעי של האמצעים, יש לפחות $\frac{n}{10}$ אמצעים גדולים או שווים ממנו (כי הסתכלנו על $\frac{n}{5}$ אמצעים), ועבור כל אחד מהאמצעים הללו יש **בדיוק** 3 איברים גדולים או שווים ממנו, לכן יש לפחות $3\cdot\frac{n}{10}$ איברים גדולים או שווים לp.

מ.ש.ל.א.☺

(ב) **צ"ל:** לנסות למצוא חסם עליון
הוכחה:
נראה לי ש$\Theta\left(n\right)$ יהיה חסם עליון על האלגוריתם של $quickselect$

מ.ש.ל.ב.☺

(ג) **צ"ל:** לכתוב נוסחא רקורסיבית ל$quickselect$ החדש

הוכחה:

נגדיר T' היעילות של $partition$, נשים לב כי

$$T'(n) = \Theta(n) + T'\left(\frac{n}{5}\right)$$

לכן נשים לב כי $T'(n) = a \cdot T'\left(\frac{n}{b}\right) + \Theta\left(n^k\right)$ עבור $a = k = 1$, $b = 5$, לכן $\boxed{q = \frac{a}{b^k} = \frac{1}{5} < 1}$,

לכן ממשפט האב מתקיים כי $\boxed{T'(n) = \Theta(n)}$,

לכן $partition$ הוא ביעילות $\Theta(n)$,

לאחר $partition$ המערך מחולק ל2 חלקים בכל אחד יש לפחות $\frac{3}{10}n$, במקרה הגרוע תמיד נבחר את המערך הגדול יותר המכיל **לכל היותר** $\frac{7}{10}n$ איברים (כי בכל צד יש לפחות $\frac{3}{10}n$)

$$\boxed{T(n) = \Theta(n) + T\left(\frac{7}{10} \cdot n\right)}$$

מ.ש.ל.ג.☺

(ד) **צ"ל:** להראות כי $T(n) = O(n)$

הוכחה:

נשים לב כי $T(n) = a \cdot T\left(\frac{n}{b}\right) + \Theta\left(n^k\right)$ עבור $a = k = 1$, $b = \frac{10}{7}$, לכן $\boxed{q = \frac{a}{b^k} = \frac{7}{10} < 1}$,

לכן ממשפט האב מתקיים כי $\boxed{T(n) = \Theta(n)}$, ומההגדרה של Θ נובע כי $\boxed{T(n) = O(n)}$

מ.ש.ל.ד.☺

(ה) **צ"ל:** מדוע זה הפתיע מתכנתים?

הוכחה:

לדעתי, זה היה מפתיע כי הפתרון הפשוט ושלרוב מצופה להיות הכי יעיל הוא למיין את המערך ולהחזיר את המקום הk.
אבל באלגוריתם הזה, ניתן לראות כי ניתן לעשות זאת **תמיד** באופן יותר יעיל מאשר למיין את כל המערך

מ.ש.ל.ה.☺