

## פתרון תרגיל מספר 5 - דאסט

שם: מיכאל גרינבאום, ת"ז: 211747639

21 באפריל 2019

1. פתרון:

(א) צ"ל: נוסחא רקורסיבית לאלגוריתם  
הוכחה:

$$T(n) = 1 + 2n + 1 + 1 + \sum_{i=1}^{n-1} T(i) = 2n + 3 + \sum_{i=1}^{n-1} T(i)$$

מ.ש.ל.א. ☺

(ב) צ"ל:  $\Theta(T(n))$   
הוכחה:

$$T(n+1) - T(n) = \left[ 2n + 3 + \sum_{i=1}^n T(i) \right] - \left[ 2n + 3 + \sum_{i=1}^{n-1} T(i) \right] = T(n) + 2$$
$$\Rightarrow \boxed{T(n+1) = 2 \cdot T(n) + 2}$$

נראה באינדוקציה כי  $2^{n-1} \leq T(n) \leq 2^{n+3}$   
בסיס:  $n = 1$

$$2^{1-1} = 1 \leq T(1) = 3 \leq 16 = 2^{1+3}$$

כלומר הטענה נכונה עבור  $n = 1$ ,  
שלב: נניח שהטענה נכונה ל  $n$  ונראה שהיא נכונה ל  $n + 1$ ,

$$T(n+1) = 2 \cdot T(n) + 2 \geq 2 \cdot T(n) \geq 2 \cdot 2^{n-1} = 2^n$$

וגם

$$T(n+1) = 2 \cdot T(n) + 3 \leq 2 \cdot 2^{n+3} + 4 = 4(2^{n+1} + 1) \leq 4 \cdot 2^{n+2} = 2^{n+4}$$

כלומר  $2^{(n+1)-1} \leq T(n+1) \leq 2^{(n+1)+3}$ , כלומר הטענה נכונה גם ל  $n + 1$ ,  
לכן  $\frac{2^n}{2} \leq T(n) \leq 8 \cdot 2^n$ ,  
כלומר  $\boxed{T(n) \in \Theta(2^n)}$  מהגדרה כי הראנו כי  $\forall n \in \mathbb{N}$  מתקיים  $\frac{1}{2} \cdot 2^n \leq T(n) \leq 8 \cdot 2^n$

מ.ש.ל.ב. ☺

2. פתרון:

(א) צ"ל: חסם עליון ל  $\mathbb{P}(X \geq \frac{n}{2})$  ומתי זה אומר משהו על ההסתברות  
הוכחה:

$$\mathbb{P}\left(X \geq \frac{n}{2}\right) \stackrel{\text{markov inequality}}{\leq} \frac{\mathbb{E}[X]}{\frac{n}{2}} \stackrel{*}{=} \frac{2 \cdot np}{n} = 2p$$

נשים לב כי  $\star$  מתקיים לפי תרגיל 3 שאלה 2 סעיף ב'.  
 עבור  $p \geq \frac{1}{2}$  נקבל כי  $\mathbb{P}(X \geq \frac{n}{2}) \leq 1$ , הדבר אינו אומר כלום כי כל הסתברות היא קטנה או שווה ל-1.  
 אך עבור  $p < \frac{1}{2}$ , אנחנו נקבל כי  $\mathbb{P}(X \geq \frac{n}{2}) \in [0, 2p]$ , לכן מפה אנחנו יכולים להסיק כי  $\mathbb{P}(X \geq \frac{n}{2}) \notin (2p, 1]$

מ.ש.ל.א. ☺

(ב) צ"ל:  $\mathbb{P}(X \geq \frac{n}{2})$   
 הוכחה:

נגדיר  $X_i = \begin{cases} i & (x_1, \dots, x_n) \text{ exactly } i \text{ of them are } H \\ 0 & \text{else} \end{cases}$ , נשים לב כי  $X = \sum_{i=0}^n X_i$

ראינו בתרגיל 3 שאלה 1 כי ההסתברות של  $(1-p)^{n-i} \cdot p^i$   $\mathbb{P}(\omega = (x_1, \dots, x_n))$ , כאשר ב  $(x_1, \dots, x_n)$  יש בדיוק  $i$  פעמים  $H$ , וגם ראינו בשאלה 1 כי מספר ה  $(x_1, \dots, x_n)$  שמכילים בדיוק  $i$  פעמים  $H$  הוא  $\binom{n}{i}$  אזי

$$\begin{aligned} \mathbb{E}\left(X \geq \frac{n}{2}\right) &= \mathbb{E}\left(\sum_{i=\lceil \frac{n}{2} \rceil}^n X_i\right) = \sum_{i=\lceil \frac{n}{2} \rceil}^n \mathbb{E}(X_i) = \sum_{i=\lceil \frac{n}{2} \rceil}^n \left[ \sum_{\omega \in \Omega} X_i(\omega) \cdot \mathbb{P}(\omega) \right] \\ &= \sum_{i=\lceil \frac{n}{2} \rceil}^n \left[ \sum_{\omega=(x_1, \dots, x_n) \text{ exactly } i \text{ of them are } H} X_i(\omega) \cdot \mathbb{P}(\omega) \right] \\ &= \sum_{i=\lceil \frac{n}{2} \rceil}^n \left[ \sum_{\omega=(x_1, \dots, x_n) \text{ exactly } i \text{ of them are } H} i \cdot (1-p)^{n-i} \cdot p^i \right] \\ &= \sum_{i=\lceil \frac{n}{2} \rceil}^n \left[ \binom{n}{i} \cdot i \cdot (1-p)^{n-i} \cdot p^i \right] \end{aligned}$$

מ.ש.ל.ב. ☺

(ג) צ"ל:  $\lim_{n \rightarrow \infty} \mathbb{P}(X_n \geq \varepsilon) = 0$   
 הוכחה:  
 נשים לב כי

$$\begin{aligned} \mathbb{E}[X_n] &= \sum_{\omega \in \Omega} X_n(\omega) \cdot \mathbb{P}(\omega) = \sum_{k \in \mathbb{N}} X_n(k) \cdot \mathbb{P}(k) = \frac{1}{n} \cdot \left[ \sum_{k=1}^n \frac{1}{k} \right] \\ &= \frac{1}{n} \cdot \left[ \sum_{k=1}^n \frac{1}{k} - \ln(n) + \ln(n) \right] = \frac{\sum_{k=1}^n \frac{1}{k} - \ln(n)}{n} + \frac{\ln(n)}{n} \end{aligned}$$

בתרגיל 4 ראינו כי  $\lim_{n \rightarrow \infty} \left[ \sum_{k=1}^n \frac{1}{k} - \ln(n) \right] = \gamma$ , לכן

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}[X_n] &= \lim_{n \rightarrow \infty} \left[ \frac{\sum_{k=1}^n \frac{1}{k} - \ln(n)}{n} + \frac{\ln(n)}{n} \right] \stackrel{*}{=} \lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n \frac{1}{k} - \ln(n)}{n} + \lim_{n \rightarrow \infty} \frac{\ln(n)}{n} \\ &\stackrel{\frac{\gamma}{\infty} \rightarrow 0}{=} 0 + \lim_{n \rightarrow \infty} \frac{\ln(n)}{n} \stackrel{L_{\infty}}{=} \lim_{n \rightarrow \infty} \frac{\frac{1}{n}}{1} = \lim_{n \rightarrow \infty} \frac{1}{n} = 0 \end{aligned}$$

נשים לב כי  $\star$  מניח שהגבולות קיימים ובהמשך ראינו כי הגבולות קיימים ושווים ל-0, עתה נשים לב כי

$$0 \leq \mathbb{P}(X_n \geq \varepsilon) \stackrel{\text{markov inequality}}{\leq} \frac{\mathbb{E}[X_n]}{\varepsilon} \rightarrow 0$$

לכן ממשפט סנדוויץ' מתקיים כי  $\lim_{n \rightarrow \infty} \mathbb{P}(X_n \geq \varepsilon) = 0$

### 3. פתרון:

(א) צ"ל: רדיקס סורט עובד

הוכחה:

נוכיח באינדוקציה (בשמורת הלולאה) כי הריצה ה- $i$  גורמת למערך להיות ממוין ביחס ל- $i$  הספרות האחרונות.  
בסיס:  $n = 1$ , המערך ממוין לפי הספרה האחרונה, ברור שהדבר מתקיים לפי נכונות אלגוריתם `stable_sort`  
שלב: נניח שהטענה נכונה ל- $n$  ונראה שהיא נכונה ל- $n + 1$ , יש עתה 2 מקרים:

- אם לשני איברים אין אותה ספרה במקום ה- $n + 1$ , אז לפי `stable_sort`, הם ממויינים כנדרש
  - אחרת, ל-2 האיברים יש אותה ספרה במקום ה- $n + 1$ , אז מהיות `stable_sort` שומר על סדר, האלגוריתם פשוט ישים אותם במיון ביחס ל- $n$  הספרות הראשונות, לכן הם יהיו ממויינים כנדרש
- כלומר בסוף הריצה, המערך יהיה ממוין ביחס לספרות עד המקום ה- $n + 1$ , כנדרש.  
 בסוף הלולאה, הכל יהיה ממוין בגלל ההנחה של שמורת הלולאה, כלומר האלגוריתם נכון!

מ.ש.ל.א. ☺

(ב) צ"ל: זמן הריצה המצופה הוא  $\mathbb{E}[n_i^2]$

הוכחה:

נשים לב כי לכל ריצה מתקיים

$$T(n) = \Theta(n) + \sum_{i=1}^k \Theta(n_i^2)$$

לכן

$$\begin{aligned} \mathbb{E}[T(n)] &= \mathbb{E}\left[\Theta(n) + \sum_{i=1}^k \Theta(n_i^2)\right] = \Theta(n) + \mathbb{E}\left[\sum_{i=1}^k \Theta(n_i^2)\right] = \Theta(n) + \sum_{i=1}^k \mathbb{E}[\Theta(n_i^2)] \\ &= \Theta(n) + \Theta\left(\sum_{i=1}^k \mathbb{E}[n_i^2]\right) \end{aligned}$$

נשים לב כי  $\sum_{i=1}^k \mathbb{E}[n_i^2] \geq n$  כי הסכום תמיד גדול שווה ל- $n$  (כי יש  $n$  איברים למיין בתוך ה-`buckets`),  
 לכן  $\Theta\left(\sum_{i=1}^k \mathbb{E}[n_i^2]\right) \geq \Theta(n)$

$$\mathbb{E}[T(n)] = \Theta(n) + \Theta\left(\sum_{i=1}^k \mathbb{E}[n_i^2]\right) \stackrel{n \in o\left(\sum_{i=1}^k \mathbb{E}[n_i^2]\right) \vee n \in \Theta\left(\sum_{i=1}^k \mathbb{E}[n_i^2]\right)}{=} \Theta\left(\sum_{i=1}^k \mathbb{E}[n_i^2]\right)$$

כלומר קיבלנו כי

$$\boxed{\mathbb{E}[T(n)] = \Theta\left(\sum_{i=1}^k \mathbb{E}[n_i^2]\right)}$$

מ.ש.ל.ב. ☺

(ג) צ"ל:  $\sum_{i=1}^k \mathbb{E}[n_i^2]$   
 הוכחה:

$$\begin{aligned}
\sum_{i=1}^k \mathbb{E}[n_i^2] &= \sum_{i=1}^k \mathbb{E} \left[ \left( \sum_{j=1}^n X_{i,j} \right)^2 \right] = \sum_{i=1}^k \mathbb{E} \left[ \left( \sum_{j_1=1}^n X_{i,j_1} \right) \cdot \left( \sum_{j_2=1}^n X_{i,j_2} \right) \right] = \sum_{i=1}^k \mathbb{E} \left[ \sum_{j_1=1}^n \sum_{j_2=1}^n (X_{i,j_1} \cdot X_{i,j_2}) \right] \\
&= \sum_{i=1}^k \mathbb{E} \left[ \sum_{j=1}^n X_{i,j}^2 \right] + \sum_{i=1}^k \mathbb{E} \left[ \sum_{j_1=1}^n \sum_{j_2=1, j_1 \neq j_2}^n (X_{i,j_1} \cdot X_{i,j_2}) \right] = \\
&= \sum_{i=1}^k \sum_{j=1}^n \mathbb{E}[X_{i,j}^2] + \sum_{i=1}^k \sum_{j_1=1}^n \sum_{j_2=1, j_1 \neq j_2}^n \mathbb{E}[(X_{i,j_1} \cdot X_{i,j_2})] \\
&\stackrel{*}{=} \sum_{i=1}^k \sum_{j=1}^n \left[ \frac{1}{k} \right] + \sum_{i=1}^k \sum_{j_1=1}^n \sum_{j_2=1, j_1 \neq j_2}^n \left[ \frac{1}{k^2} \right] = k \cdot n \cdot \frac{1}{k} + k \cdot \frac{1}{k^2} \cdot n \cdot (n-1) \\
&= n + \frac{n^2 - n}{k} = \frac{n^2 + nk - n}{k}
\end{aligned}$$

נשים לב כי  $\star$  מתקיים כי

$$\begin{aligned}
\mathbb{E}[X_{i,j}^2] &= 1^2 \cdot \frac{1}{k} + 0^2 \cdot (1-k) = \frac{1}{k} \\
\mathbb{E}[(X_{i,j_1} \cdot X_{i,j_2})] &= 1^2 \cdot \frac{1}{k} \cdot \frac{1}{k} = \frac{1}{k^2}
\end{aligned}$$

לכן

$$\boxed{\sum_{i=1}^k \mathbb{E}[n_i^2] = \frac{n^2 + nk - n}{k}}$$

נראה כי

$$\boxed{T(n) = \Theta\left(\frac{n^2 + nk - n}{k}\right) = \Theta\left(\frac{n^2}{k} + n\right)}$$

**הערה:** לא התייחסו לאיחוד המערכים, בהנחה ואיחוד  $k$  מערכים לוקח  $\Theta(k)$  אז  $T(n) = \Theta\left(\frac{n^2}{k} + n + k\right)$

מ.ש.ל.ג.⊙

4. פתרון:

(א) צ"ל:  $L(a) < L(b)$

**הוכחה:**

מהיות  $a \neq a_h$ , קיים עלה של  $t$  שגובהו קטן מ  $h$  (אחרת  $a$  היה עץ מלא והיינו מקבלים  $a = a_h$ ) נגדיר  $b$  להיות העץ  $a$  רק של  $t$  יש 2 עלים, אזי  $L(a) < L(a) + 1 = L(b)$  (נשים לב שגובה העלים שהוספנו הם קטנים שווים  $h$  כי הגובה של  $t$  קטן מ  $h$  ולכן  $b \in A_h$ )

מ.ש.ל.א.⊙

(ב) צ"ל:  $L(a_h) = \max\{L(a) \mid a \in A_h\}$

**הוכחה:**

נניח בשלילה כי  $L(a_h) \neq \max\{L(a) \mid a \in A_h\}$ , נסמן את העץ המקסימלי ב  $a$ , מהיות  $a \neq a_h$ , מהסעיף הקודם קיים  $b \in A_h$  כך ש

$$L(b) > L(a) = \max\{L(a) \mid a \in A_h\} \geq L(b)$$

סתירה כי קיבלנו  $L(b) > L(b)$ , לכן  $L(a_h) = \max\{L(a) \mid a \in A_h\}$

מ.ש.ל.ב. ☺

$$L(a_h) = 2^h \text{ (ג) צ"ל: } L(a_h) = 2^h$$

הוכחה:

נוכח את הטענה באינדוקציה:

בסיס:  $h = 1$ , יש 2 עלים בעץ לכן  $L(a_1) = 2 = 2^1$ , כלומר הטענה נכונה ל  $h = 1$ ,

שלב: נניח שהטענה נכונה ל  $h$  ונראה שהיא נכונה ל  $h + 1$ ,

מההגדרה של עץ מלא לכל ענף בגובה  $h$  יש 2 ילדים והם עלים היחידים, לכן

$$L(a_{h+1}) = 2 \cdot L(a_h) \stackrel{\text{induction}}{=} 2 \cdot 2^h = 2^{h+1}$$

כלומר הטענה נכונה ל  $h + 1$ , לכן מתקיים כי  $L(a_h) = 2^h$   $\forall h \in \mathbb{N}$

מ.ש.ל.ג. ☺

5. פתרון:

$$\Theta(T(n)) \text{ (א) צ"ל: } \Theta(T(n))$$

הוכחה:

מהיות  $Merge$  הוא ביעילות  $\Theta(|A| + |B|)$  קיימים  $c_1 > 0, c_2 > 0$  כך ש

$$c_2 \cdot (|A| + |B|) \leq Merge \leq c_1 \cdot (|A| + |B|)$$

עתה נשים לב כיניתן להראות באינדוקציה ש  $|B_{i-1}| = \frac{i \cdot n}{k}$ , לכן

$$\begin{aligned} T(n) &= 2 + k + \sum_{i=1}^k Merge(B_{i-1}, A_i) \leq 2 + k + \sum_{i=1}^k c_1 \cdot (|B_{i-1}| + |A_i|) \\ &= 2 + k + c_1 \cdot \sum_{i=1}^k \left( \frac{n}{k} + \frac{n \cdot i}{k} \right) = 2 + k + \frac{c_1 \cdot n}{k} \cdot \sum_{i=1}^k (i + 1) \\ &= 2 + k + \frac{c_1 \cdot n}{k} \cdot \sum_{i=1}^k (i + 1) \leq 2 + k + \frac{c_1 \cdot n}{k} \cdot \sum_{i=1}^k i = 2 + n + \frac{c_1 \cdot n}{k} \cdot \frac{k(k+1)}{2} \\ &= 2 + k + c_1 \cdot n \cdot \frac{(k+1)}{2} = \frac{c_1}{2} \cdot n \cdot k + \frac{c_1}{2} \cdot n + k + 2 \stackrel{\frac{c_1}{2} \cdot n + k + 2 = o(n \cdot k)}{\in} O(n \cdot k) \end{aligned}$$

לכן  $T(n) = O(n \cdot k)$  עתה נשים לב כי

$$\begin{aligned} T(n) &= 2 + k + \sum_{i=1}^k Merge(B_{i-1}, A_i) \geq 2 + k + \sum_{i=1}^k c_2 \cdot (|B_{i-1}| + |A_i|) \\ &= 2 + k + c_2 \cdot \sum_{i=1}^k \left( \frac{n}{k} + \frac{n \cdot i}{k} \right) = 2 + k + \frac{c_2 \cdot n}{k} \cdot \sum_{i=1}^k (i + 1) \\ &= 2 + k + \frac{c_2 \cdot n}{k} \cdot \sum_{i=1}^k (i + 1) \geq 2 + k + \frac{c_1 \cdot n}{k} \cdot \sum_{i=1}^{k-1} i = 2 + n + \frac{c_1 \cdot n}{k} \cdot \frac{k(k-1)}{2} \\ &= 2 + k + c_1 \cdot n \cdot \frac{(k-1)}{2} = \frac{c_1}{2} \cdot n \cdot k - \frac{c_1}{2} \cdot n + k + 2 \stackrel{k - \frac{c_1}{2} \cdot n + 2 = o(n \cdot k)}{\in} \Omega(n \cdot k) \end{aligned}$$

לכן  $T(n) = \Theta(n \cdot k)$   $T(n) = \Omega(n \cdot k)$  לכן

מ.ש.ל.א. ☺

(ב) צ"ל: אלגוריתם ביעילות  $O(n \cdot \log(k))$

הוכחה:

במקום לעשות איחוד של המערכים אחד אחרי השני, נחבר את הראשון לשני, את השלישי לרביעי, החמישי לשישי וכך הלאה, ונאחד ביניהם באותו אופן בצורה רקורסיבית. נשים לב כי אנחנו ביצענו לכל היותר  $\log(k)$  איחודים כשהיעילות של איחוד היא  $O(n)$ , לכן האלגוריתם הזה ירוץ ביעילות יותר טובה (או טובה כמו) מ- $O(n \cdot \log(k))$

מ.ש.ל.ב. ☺

(ג) צ"ל: כל אלגוריתם ירוץ ביעילות  $\Omega(n \cdot \log(k))$

הוכחה:

נניח בשלילה כי האלגוריתם רץ ביעילות יותר טובה מ- $n \cdot \log(k)$ , ונגדיר את היעילות שלו ב- $\Theta(n \cdot f(k))$ , (ראינו לאיחוד מערכים נדרש לפחות  $n$  פעולות).

לכן ניתן לאלגוריתם  $n$  מערכים בגודל 1 ונקבל כי האלגוריתם מחזיר מערך ממוין ביעילות  $\Theta(n \cdot f(n))$ .  
ראינו בהרצאה שכל אלגוריתם תלוי השוואות המחזיר מערך ממוין הוא לפחות ביעילות  $\Omega(n \cdot \log(n))$ ,  
לכן קיבלנו סתירה להנחה שלנו כי ניתן לבצע את האלגוריתם ביעילות יותר טובה מ- $n \cdot \log(k)$ ,  
לכן כל אלגוריתם ירוץ ביעילות  $\Omega(n \cdot \log(k))$ , כנדרש

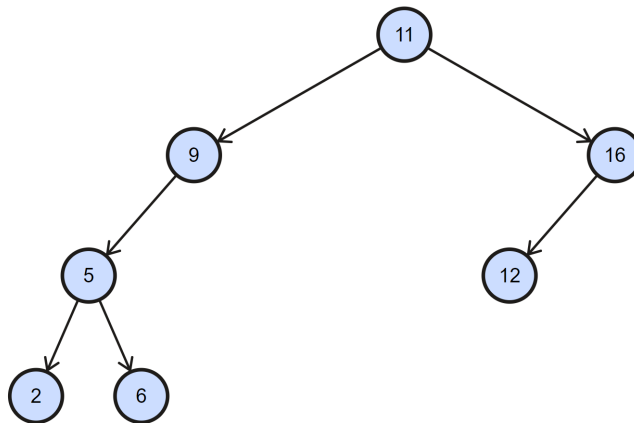
מ.ש.ל.ג. ☺

6. פתרון:

(א) צ"ל: העץ שנוצר מהוספה של האיברים

הוכחה:

ההרצה של האלגוריתם יוצרת את העץ הבא:



מ.ש.ל.א. ☺

(ב) צ"ל: מה *inorder* עושה?

הוכחה:

ההרצה של *Inorder* על העץ שלנו מדפיס 2, 5, 6, 9, 11, 12, 16.  
האלגוריתם *Inorder* מדפיס את העץ באופן ממוין מהקטן לגדול.

מ.ש.ל.ב. ☺

(ג) פתרון:

i. צ"ל: לאחר ש-*inorder* יוצא מתת עץ, האלגוריתם לא חוזר לשם

הוכחה:

נשים לב כי האלגוריתם מתנהג בדיוק כמו האלגוריתם הבא:

```
def Inorder(tree):
    if tree == null:
        return
    Inorder(tree.left)
    print(tree.value)
    Inorder(tree.right)
```

מכיוון שאנחנו הולכים למינימלי שראינו שהוא הכי שמאלי, לאחר מכן חוזרים לאב שלו ולבסוף הולכים שמאלה, בדיוק כמו שנתון *Inorder* שנתון עובד. לכן, נשים לב שכאשר אנחנו יוצאים מתת עץ, איננו נחזר אליו אחרת נדפיס את אותו הערך פעמיים בסתירה לכך שהאלגוריתם אכן מדפיס כל איבר בסדר הנכון **פעם אחת בדיוק**.

מ.ש.ל.ג.1. ☺

ii. **צ"ל:** מגיעים לכל עלה לכל היותר פעמיים

**הוכחה:**

נשים לב כי האלגוריתם מתנהג בדיוק כמו האלגוריתם הבא:

```
def Inorder(tree):
    if tree == null:
        return
    Inorder(tree.left)
    print(tree.value)
    Inorder(tree.right)
```

מהנימוק שניתן בסעיף הקודם. יתרה מכך, הפעמים היחידות שנגיע לאות העלה תהיה לראשונה כשנגיע ל-*tree.left*, אחר כך ל-*tree.right* ולבסוף ל-*tree* עצמו, כלומר עברנו דרכו לכל היותר פעמיים ויצאנו לאחר *tree.right*

מ.ש.ל.ג.2. ☺

iii. **צ"ל:** יעילות האלגוריתם

**הוכחה:**

נראה כי האלגוריתם רץ ב- $\Theta(n)$ ,

תחילה נשים לב שהאלגוריתם רץ על כל עלה לפחות פעם אחת, לכן יעילות האלגוריתם היא לפחות  $n$ ,

כלומר  $T(n) = \Omega(n)$ ,

בנוסף לכך, האלגוריתם רץ פחות מ-4 פעמים על כל עלה ומבצע לכל היותר 5 פעולות על כל עלה, לכן  $T(n) \leq 20n$ ,

לכן  $T(n) = O(n)$ , לכן  $T(n) = \Theta(n)$

מ.ש.ל.ג.3. ☺

(ד) **צ"ל:** מדוע זה אפשרי שהאלגוריתם רץ בפחות מ- $n \cdot \log(n)$

**הוכחה:**

נשים לב שעל אף שאין לנו הנחות על גודל האיברים בעץ, יש לנו הנחה שהעץ מסודר בדרך מיוחדת לבחירתנו.

לכן, עבור הסידור הספציפי הזה, ניתן לרוץ ביעילות  $\Theta(n)$ , כמו שניתן לרוץ ב- $\Theta(n)$  על מערך ממזין (מההנחה והוא כבר ממוין).

מ.ש.ל.ד. ☺