

## בתוכנית

### פרק 12 בספר הלימוד

- נכיר את ה- ADT "מילון"
- נכיר את מבנה הנתונים "עץ חיפוש בינארי" למימוש מילון, ועוד כמה תכונות ואלגוריתמים שימושיים בעצים

2

# מבנה נתונים ומבוא לאלגוריתמים

## נושא 8

### עץ חיפוש ביןaries Binary search trees

אמיר רובינשטיין

## מבנה נתונים למילון

- בימוש של מילון בכל אחד מבני הנתונים שלמדו עד כה, לפחות אחת מהפעולות תבוצע בזמן ליניארי במספר האיברים.
- נראה כתם שימוש אפשרי למילון, באמצעות מבנה הנתונים עץ חיפוש ביןארי.  
במימוש כל הפעולות יבוצעו בסיבוכיות של  $(n \log n)$ .
- בשיעור הבא נראה שיכלול של עץ חיפוש ביןaries - עץ AVL - שմבטיח סיבוכיות של  $(n \log n)$  גם במקרהorst.

4

## מילון (dictionary)

מילון הוא ADT המוגדר ע"י פעולות הבאות:

- $x, S$  – הכנסת האיבר אליו מצביע  $x$  למבנה  $S$
- $x, S$  – מחיקת האיבר אליו מצביע  $x$  מהמבנה  $S$
- $S, k$  – חיפוש איבר שמספרתו  $k$  והחזרתו, אם קיים בו  $S$

ולפעמים גם פעולות חיפוש נוספת, כמו:

- $S$  – החזרת איבר בעל מפתח מינימלי
- $S$  – החזרת איבר בעל מפתח מקסימלי
- $x, S$  – החזרת עוקב של  $x$  ב-  $S$  (איבר בעל מפתח מינימלי שגדל מזה של  $x$ )
- $x, S$  – החזרת קדום של  $x$  ב-  $S$  (איבר בעל מפתח מקסימלי שקטן מזה של  $x$ )

4 הפעולות האחרונות מניחות שיש סדר על מפתחות האיברים.

מילון שימושים רבים ומגוונים (למשל?)

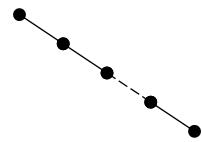
3

## גובה של עץ בינארי

נסמן את מספר הצלטמים בעץ בינארי ב-  $a$ , ואת גובהו ב-  $h$ .

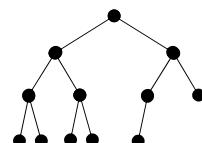
מהו הגובה המינימלי ומהו הגובה המקסימלי של עץ בינארי, כתלות ב-  $a$ ?

$$h = n-1 = \Theta(n)$$



מקסימלי: שרשרת או "זיג-זג"

$$h = \lfloor \log_2 n \rfloor = \Theta(\log n)$$



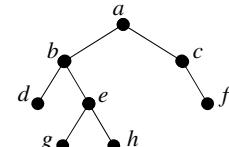
מינימלי: עץ (כמעט) שלם

6

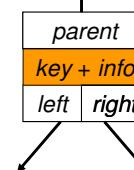
## ~~עץ חיפוש בינארי - תזכורת~~

### מונחים - תזכורת

- צומת
- קשת
- שורש
- בן השמאלי, בן ימני, אב, אב-קדמון, צאצא
- תת-עץ שמאל, תת-עץ ימני
- עליה, צומת פנימי
- עומק של צומת
- גובה של צומת, של עץ
- שימוש במערך, מימוש באמצעות רשומות עם מצביעים:



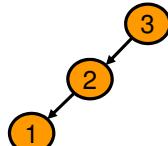
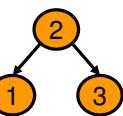
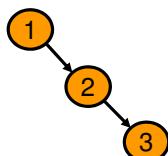
מבנה של רשומה (צומת):



5

## עץ חיפוש בינארי

בහינתן קבוצת מפתחות, האם קיימן עץ חיפוש בינארי יחיד המכיל מפתחות אלו?



עוד?

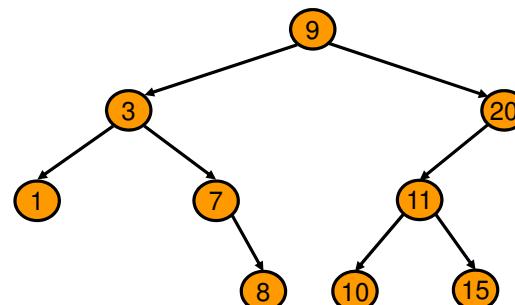
8

## עץ חיפוש בינארי

עץ חיפוש בינארי (binary search tree) הוא עץ בינארי עם התכונה הבאה:

בහינתן צומת עם מפתח  $k$ :

- צלטמים בתת העץ הימני שלו בעלי מפתחות  $\leq k$
- צלטמים בתת העץ השמאלי שלו בעלי מפתחות  $\geq k$ .



בහינתן קבוצת מפתחות, האם קיימן עץ חיפוש בינארי יחיד המכיל מפתחות אלו?

7

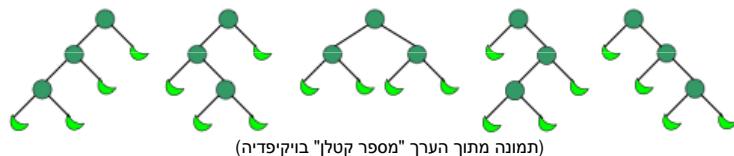
## מספרי קטלן (Catalan numbers)

מספרי קטלן מיצגים גם את מספר הדרכים השונות לשים סוגרים על  $1+n$  גורמים שונים.

לדוגמה, כאשר  $n=3$  יש 5 דרכים שונות לשים סוגרים על 4 גורמים:

$d$        $(ab)c$        $(ab)(cd)$        $a(bc)d$        $a(b(cd))$

ישנה פונקציה  $\text{ch}(n)$  ועל בין קבוצת הביטויים הנ"ל לבין קבוצת העצים הבינאריים בעלי  $n$  צמתים:



10

## כמה עצים בינאריים שונים בעלי $n$ צמתים קיימים?

נסמן את התשובה ב-  $b_n$ .

נוסחה רקורסיבית עבור  $b_n$ :

עבור  $0 \leq n$ : יש רק עץ אחד ללא צמתים – העץ הריק. לכן,  $b_0 = 1$ .

עבור  $1 \leq n$ : העץ מורכב מצומת אחד בשורש,

$k$  צמתים בתת-עץ השמאלי ( $k = 0, 1, \dots, n-1$ )

$n-k-1$  צמתים בתת-עץ הימני.

לכן:

$$b_n = \sum_{k=0}^{n-1} b_k \cdot b_{n-k-1}$$

$$b_n = \frac{1}{n+1} \binom{2n}{n}$$

$$= \frac{4^n}{\sqrt{\pi n}^{3/2}} (1 + O(1/n))$$

9

$b_n$  נקראים **מספרי קטלן** (Catalan numbers).

## חיפוש

Tree-Search( $x, k$ )

1. **while**  $x \neq \text{nil}$  **and**  $k \neq \text{key}[x]$
2.   **if**  $k < \text{key}[x]$
3.      $x \leftarrow \text{left}[x]$
4.   **else**  $x \leftarrow \text{right}[x]$
5. **return**  $x$

נתן גם לכתוב גרסה איטרטיבית:

סיבוכיות זמן:  $\Theta(h)$

סיבוכיות זיכרון נוספת:  $\Theta(1)$

שאלה: מהו סיבוכיות הזמן במקרה הטוב?

12

## חיפוש

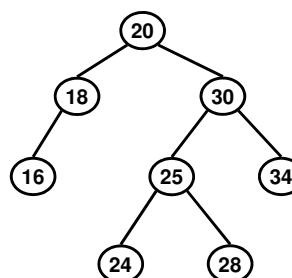
כיצד נחשף מפתח בעץ חיפוש בינארי?

Tree-Search( $x, k$ )

1. **if**  $x = \text{nil}$  **or**  $k = \text{key}[x]$
2.   **return**  $x$
3. **if**  $k < \text{key}[x]$
4.   **return** Tree-Search( $\text{left}[x], k$ )
5. **else return** Tree-Search( $\text{right}[x], k$ )

בקיראה הראשונה  
 $x$  הוא שורש העץ.

מהו מסלול החיפוש של 26 בעץ הבא?



מה סיבוכיות הזמן והזיכרון הנוסף?

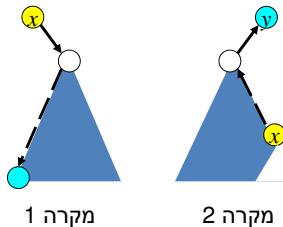
סיבוכיות זמן:  $\Theta(h)$

סיבוכיות זיכרון נוספת:  $\Theta(h)$

11

## שאלות נוספת – עוקב וקדם

- הזרת האיבר העוקב לאיבר נתון  $x$



**Tree-Successor( $x$ )**

- if**  $right[x] \neq \text{nil}$
- return** Tree-Min( $right[x]$ )
- $y \leftarrow parent[x]$
- while**  $y \neq \text{nil}$  **and**  $x = right[y]$
- $x \leftarrow y$
- $y \leftarrow parent[y]$
- return**  $y$

14

מקרה 2:

מימוש הזרת הקודם - סימטרי.  
סיבוכיות זמן:  $\Theta(h)$ .

## שאלות נוספות – מינימום ומקסימום

חיפוש משטייך לסוג של פעולות הנקראות **שאלות**.  
פעולות אלו אין שינוי של מבנה הנתונים, אלא רק שילוף מידע ממנה.

שאלות נוספות:

- הזרת איבר בעל מפתח מינימלי:  
(הזרת המינימום - סימטרי)

סיבוכיות זמן:  $\Theta(h)$

ובמקרה הטוב?

Tree-Min( $x$ )

- while**  $left[x] \neq \text{nil}$
- $x \leftarrow left[x]$
- return**  $x$

13

## הכנסה

האלגוריתם:

**Tree-Insert( $T, z$ )**

- $y \leftarrow \text{nil}$
- $x \leftarrow root[T]$
- while**  $x \neq \text{nil}$
- $y \leftarrow x$
- if**  $key[z] < key[x]$
- $x \leftarrow left[x]$
- else**  $x \leftarrow right[x]$
- $parent[z] \leftarrow y$
- if**  $y = \text{nil}$
- $root[T] \leftarrow z$
- else if**  $key[z] < key[y]$
- $left[y] \leftarrow z$
- else**  $right[y] \leftarrow z$

16

רוצים ש- $z$  מכיל שדות מצביים לבן שמאל וימני ולאב.

- שורות 1-7:  
מחפשים אב  $y$  ש"מוכן לאמץ" את  $z$   
(כך שתישמר תכונת עץ החיפוש הבינארי)

- שורות 8-13:  
מציבים את  $z$  כבן (שמאל או ימני) של  $y$

- שורות 9-10:  
טיפול במקרה הקצה –  $z$  הוכנס לעץ ריק

סיבוכיות זמן:  $\Theta(h)$ .

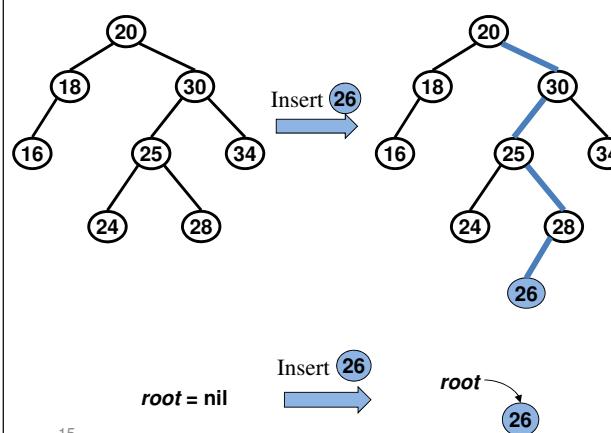
## הכנסה

סוג אחר של פעולות **הכנסה**, המשנות את מבנה הנתונים (כמו הכנסה, מחיקה, שינוי).

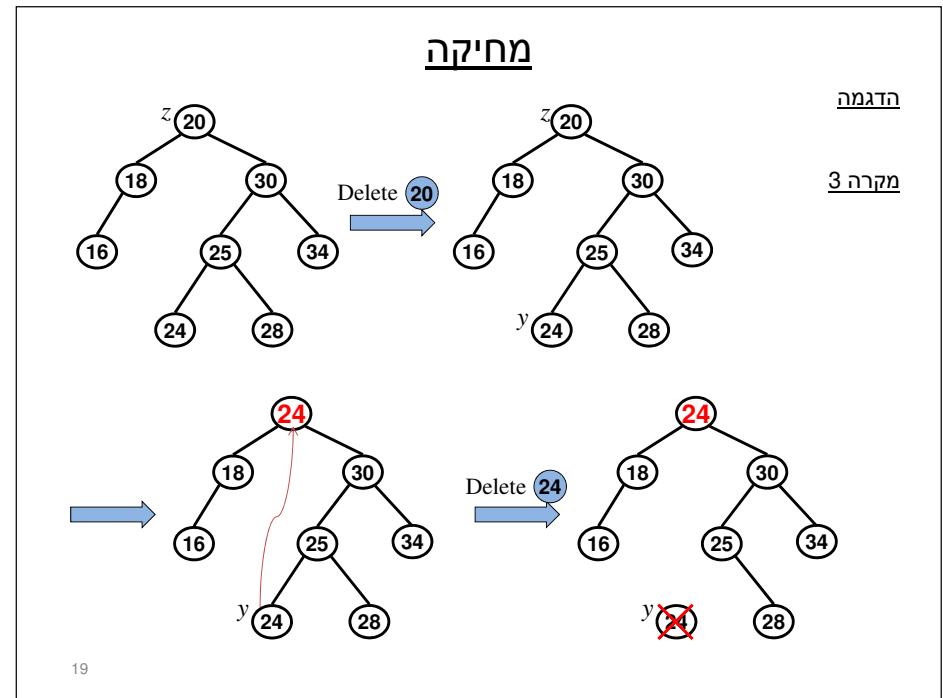
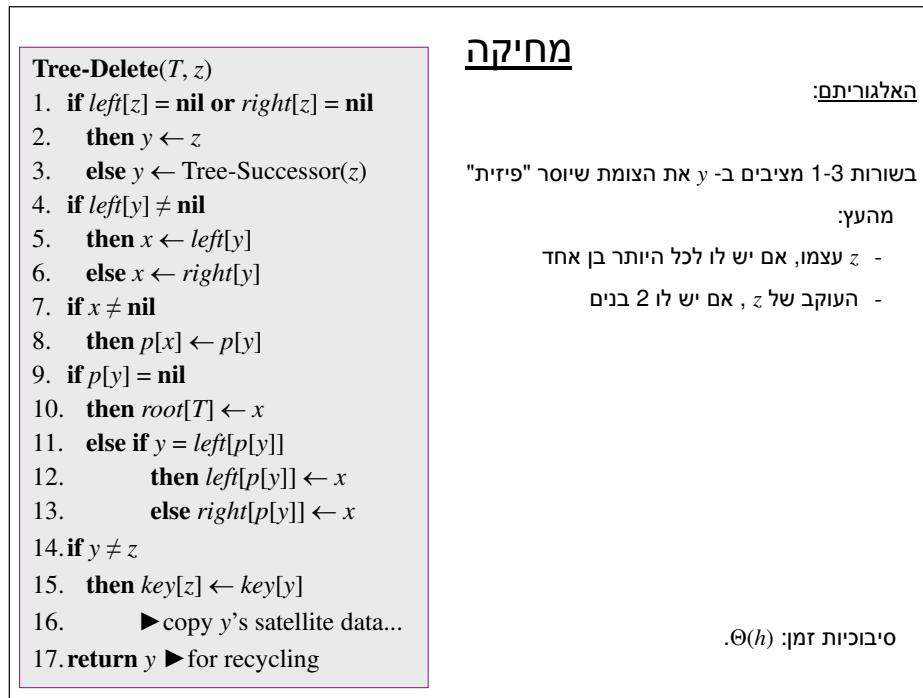
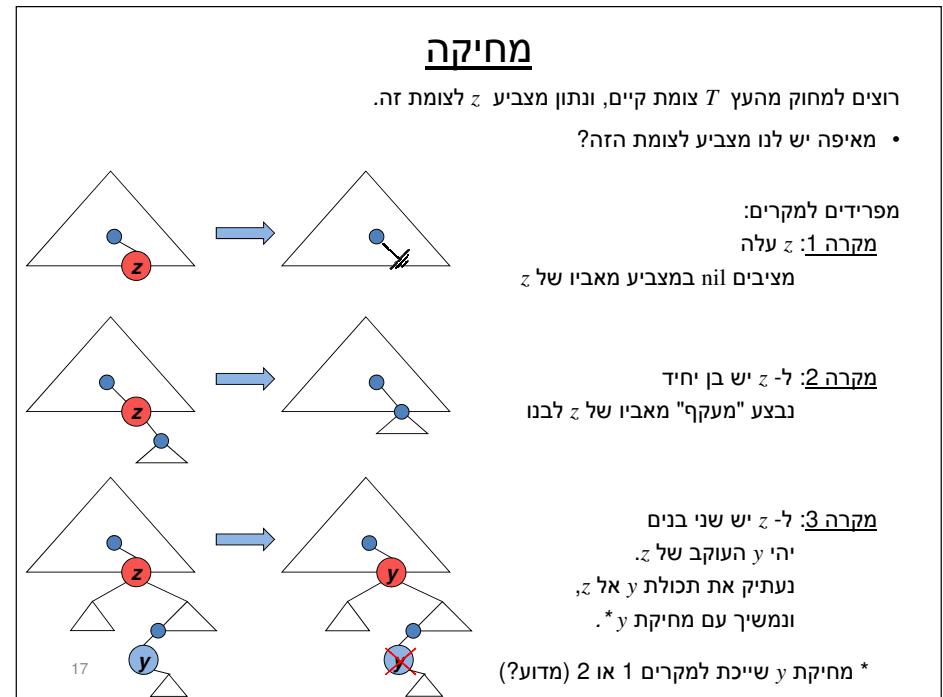
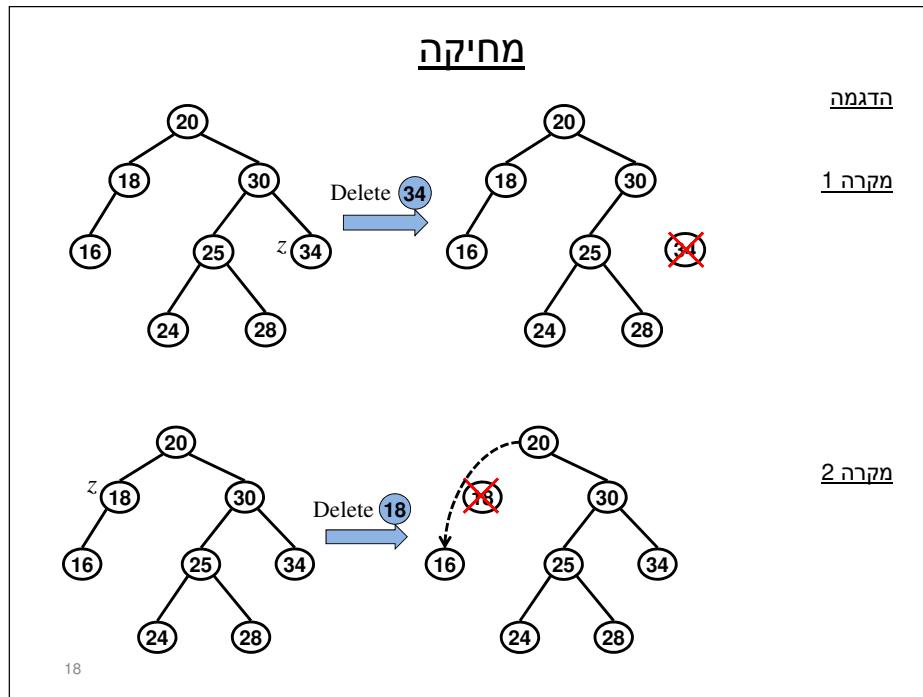
הכנסה

מחפשים את מקום הכנסה המתאים (בדומה לחיפוש רגיל) ו"תולים" את האיבר החדש כעליה.

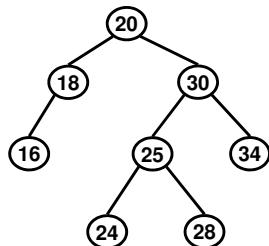
הדוגמה



הכנסה לעץ ריק:



## סיוורים בעצים



תרגיל: רישמו את סדר הביקור בצלמיים בכל אחד מהסיוורים.

• סיוור pre-order:

• סיוור in-order:

• סיוור post-order:

22

## סיוורים בעצים

לעתים אנו מעוניינים לעבור באופן שיטתי על כל צמתים של עץ בינהר למטרת כלשהי.

פעולה זאת נקראת סיוור בעץ או סריקה של העץ.

נכיר 3 שיטות לסריקת עץ בינהר, הנבדלות זו מזו בסדר בו נסרקים הצלמיים.

סיוור תחيلي (pre-order)

- מבקרים בשורש

- מסיירים רקורסיבית בתת העץ השמאלי

- מסיירים רקורסיבית בתת העץ הימני

סיוור תוכי (in-order)

הביקור בשורש געשה בין הסיוור בתת העץ השמאלי שלו לבין הסיוור בתת העץ הימני שלו.

סיוור סופי (post-order)

הביקור בשורש געשה אחרי הסיוורים בתתי העצים השמאלי והימני שלו.

21

## סיוורים בעצים

### סיבוכיות זמן

טענה: כל אחד משלושת הסיוורים הנ"ל רץ על עץ בעל  $n$  צמתים בזמן ( $n$ ) $\Theta$ .

אינטואיציה: עוברים דרך כל צומת 3 פעמים: בדרך לבנו השמאלי, בדרך לבנו הימני, ובדרך לאבוי.

$$T(n) = 2T(n/2) + 1 = \Theta(n)$$

הוכחה שגוייה: נוסחת הנסיגה המתאימה היא:

מה הטעות?

הוכחה: נסמן את מספר הצלמיים בתת-העץ השמאלי של השורש ב-  $k$ .

נניח שביקור בצלמת דרוש  $d > k$  פועלות.

$$T(0) = 1$$

$$T(n) = T(k) + T(n-k-1) + d$$

נוכיח בשיטת ההצבה כי  $T(n) = (d+1)n + 1$

$$T(0) = (d+1) \cdot 0 + 1 = 1$$

בסיום:

$$T(n) = (d+1)k + 1 + (d+1)(n-k-1) + 1 + d = (d+1)n + 1$$

צעד:

24

## סיוורים בעצים

### טענה:

סיוור order-in בעץ חיפוש בינהר, שבו Visit מדפסה את מפתח הצומת, מניבה סדרה ממוגנת (של מפתחות העץ).

### הוכחה:

באינדוקציה על מספר הצלמיים בעץ.

בבסיס: עבור עץ עם צומת אחד הטענה בודאי נכונה.

צעד: עבור עץ עם יותר מצומת אחד, יודפסו קודם כל הצלמיים בתת העץ השמאלי של השורש, אח"כ השורש ורף לאחר מכן כל הצלמיים בתת העץ הימני שלו. לפי הנקחת האינדוקציה, הסיוור בכל מעתה העצים מניב סדרה ממוגנת, ולפי תכונת עץ החיפוש הסדרה יכולה ממוגנת.

23



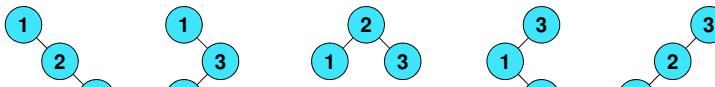
## גובה ממוצע

ראינו קודם שהגובה הממוצע של עץ בינארי הוא  $1 - \frac{1}{n}$ , ואילו הגובה המינימלי הוא `[logn]`.

שאלה: רצים להכנס לעץ חיפוש ביןארי את המפתחות 1,2,3.

- כמה סדרי הכנסה שונים אפשריים?
- אילו מהם יניבו עץ בגובה מינימלי?
- אילו מהם יניבו עץ בגובה ממוצע?

$<1,2,3>$      $<1,3,2>$      $<2,1,3>$  and  $<2,3,1>$      $<3,1,2>$      $<3,2,1>$



גובה ממוצע הוא ממוצע האגדהים של העצים המתאפשרים בסדרי הכנסה השונים (בדוגמה:  $\frac{1^2 + 2^2}{3}$ ).

30

## שימושים של סוררים - ביטוי postfix

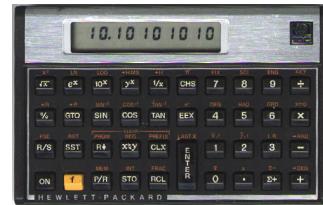
האלגוריתם לחישוב ביטוי postfix

1. התחל עם מחסנית ריקה.

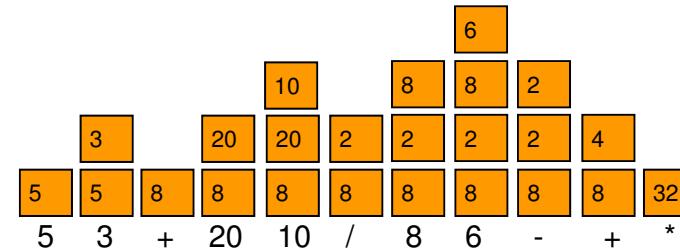
2. עבר על הביטוי משמאלי לימין.

3. אם האיבר הבא הוא אופרנד - הכנס אותו למחסנית.

4. אם הוא פעולה – הפעל את הפעולה על שני האיברים בראש המחסנית והכנס את התוצאה למחסנית.



HP-10c programmable scientific, 1980's.  
Cost: \$80



:postfix

29

## שאלות חזרה

1. בקרו בקישור הבא, המציג סימולציה מצוינת של עץ חיפוש ביןארים (ועוד סוגים אחרים של עץ חיפוש):

<http://people.ksp.sk/~kuko/bak/index.html>

היעזרו בסימולציה כדי להבין את האלגוריתמים השונים על עץ חיפוש ביןארי.

אתרי הדגמות נוספים, שעשוים לשיער:

א. <http://nova.umuc.edu/~jarc/idsv/> (יש לבחור Search Trees בתפריט)

ב. <http://www.site.uottawa.ca/~stan/csi2514/applets/avl/BT.html> (יש לבטל את הסימון ליד AVL).

2. צירואת כל העצים האפשריים המתאפשרים בעץ הכנסת האיברים 1,2,3,4 לעץ חיפוש ביןארי.

3. הציעו אלגוריתם רקורסיבי לחישוב מספר הצמתים בעץ ביןארי. מהי סיבוכיות הזמן והמקום שלו?

## גובה ממוצע

משפט:

גובה ממוצע של עץ חיפוש ביןארי בעל  $n$  צמתים הוא  $(\log n)/\Theta$ , כאשר הממוצע הוא על פני כל סדרי הכנסה השונים (הכנסות בלבד, ללא מחיקות או פעולות אחרות).

הוכחה בספר הלימוד, עמודים 225-223.

ניסוח אחר: כאשר סדר הכנסת האיברים הוא אקראי, תוחלת הגובה היא לוגריתמית במספר הצמתים.

32

31

## תרגילים

34

### תרגילים נוספים

- נתון עץ בינארי בעל  $n$  צמתים, שbulletן צומת שלו יש מספר חיובי. הצעו אלגוריתם רקורסיבי לחישוב מסלול כבד ביותר מהשורש לעלה – מסלול מהשורש לעלה, שסכום ערכי הצמתים לאורכו הוא מוקטני. הצעו גם דרך להדפיס את צמחי מסלול זה. מה היה משתנה אילו המספרים היו לא בהכרח חיוביים?
- סטודנט הפעיל סירורים  $-inorder$  ו-  $-preorder$  על עץ בינארי שהיה ברשותו. כמובן זן מה, הוא גילה לתזהמתו שאיבד את העץ, אך יש עדין בידיו את תוצאות הסירורים. עזרו לסטודנט לשחזר את העץ. מה סיבות הפתרון במקרה הגראע / הטוב?

*inorder* : 2 6 4 7 1 3 8 5 9 10

*preorder* : 1 2 4 6 7 3 5 8 9 10

- נתן למניין קבוצה נתונה של  $n$  מספרים באופן הבא:  
 - תחיליה בונים עץ חיפוש בינארי המכיל מספרים אלו, ע"י הכנסת האיברים זה אחר זה לעץ.  
 - לאחר מכן מדפיסים את המספרים בסדר המתkeletal ע"י סירור *inorder* של העץ.  
 מהו זמן הריצה במקרה הגראע, במקרה הטוב, ובמוצע, של אלגוריתם מין זה?

36

## תשובות לשאלות חזרה

3. האלגוריתם:

```
Count(x)
1. if  $x = \text{NIL}$ 
2.   return 0
3.   return 1 + Count( $\text{left}[x]$ ) + Count( $\text{right}[x]$ )
```

סיבות זמן ליניארית במספר הצמתים.  
סיבות זיכרון ליניארית בגובה העץ.

33

### תרגילים מומלצים מהספר

#### פרק 12

12.1-1

12.1-2

12.1-3

12.1-4

12.1-5

12.2-1

12.2-2

12.2-3

12.2-4

12.2-5

12.2-6

12.2-7

\*12.2-8

12.2-9

12.3-1

12.3-2

12.3-3

\*12.3-4

12.3-5

12.3-6

בעה 12-1  
בעה 12-2 (שיר לנושא של מחירות)

35

#### תרגילים נוספים

4. הוכיחו כי לא קיים אלגוריתם לבניית עץ חיפוש בינארי מרשים נתונה של  $n$  איברים, שזמן ריצתו  $(n \log n)$ .

#### 5. ואראיציה של תרגיל 12.2-7 מספר הלימוד

נתן למשר סריקה תΟ(1-in-order) של עץ חיפוש בינארי בעלי  $n$  צמתים ע"י מציאת המינימום, וachs"כ מציאת העקב של הצומת הנוכחי 1-a פעמיים.

להלן חישוב חסם הדוק שגוי בזמן הריצה במרקחה הגראע:

- גובה העץ הוא  $\Theta(n)$ .

- כל אחת מפעולות Tree-Successor ריצה בזמן  $\Theta(h)$ .

- לכן סה"כ זמן הריצה הוא  $\Theta(n^2)$ .

הסבירו היכן הטעות ומדווע חישוב זה מתאים לחסם עליון בלבד (רמז: איפה צריך להחליף  $\Theta$  ב- $O$ ?).

הסבירו מדוע למעשה מעשה זמן הריצה של אלגוריתם זה הוא  $\Theta(n)$ .

37

```
Heavy-Weight( $x$ )
1. if  $x = \text{nil}$ 
2.   return 0
3.  $l \leftarrow \text{Heavy-Weight}(\text{left}[x])$ 
4.  $r \leftarrow \text{Heavy-Weight}(\text{right}[x])$ 
5. return  $\text{key}[x] + \max\{l, r\}$ 
```

תנו דוגמה שבה הפתרון אינו נכון, אם מותרים גם מספרים שליליים בצתמים.  
מה צריך לשנות בפתרון במרקחה הגראע?

38

#### פתרון 1

נניח בשלילה שקיים צה אלגוריתם.  
אז ניתן למיין  $n$  איברים בזמן  $(n \log n)$  באופן הבא:

- בניית העץ בעדרת האלגוריתם הב"ל, בזמן  $(n \log n)$ .

- בצע סירור inorder בעץ להדפסת איברי, בזמן  $(n \Theta)$ .

סה"כ סיבוכיות הזמן היא  $(n \log n)$ , וזה בסתיו לחסם התחתון לביעית מינ' ההשוואות שהוכחנו.

#### פתרון 4

אמנם, מציאת עוקב ריצה במרקחה הגראע -  $(h \Theta)$ , אבל באופן סריקה המוצג בשאלת רק חלק מהפעמים מתרחש המקרה הגראע, ופעמים רבות מציאת העוקב תיקח זמן קצר יותר  
(תנו דוגמה למצב בו מציאת העוקב של צומת ריצה בזמן קבוע). ניתן אם כן לתקן ולומר שככל אחת מפעולות העוקב ריצה בזמן  $(O(n))$ , ושהחסם העלין בזמן הריצה הכלול הוא  $(n^2 O)$ .

חסם הדוק:

נשים לב שסדר המעבר על קשתות וצמתים העז זהה לחלוון זהה שבסריקה in-order.

כלומר עוביים בכל צומת ובכל קשת אותו מספר פעמים ובאותו סדר בדיק.

ידעו שסריקה inorder בעץ ריצה בזמן ליניארי במספר צמתיו.

39