

מבני נתונים - תרגול 9

תרגול על עצים

גלעד אשרוב

20 במאי 2012

משפחה של עצים תיקרא מאוזנת אם כל עץ הוא בעל גובה $O(\log n)$, כאשר n הוא מספר הצמתים בעץ. לכל תכונה בהמשך, קבע האם משפחה של עצים המקיימים את התכונה היא מאוזנת. אם התשובה היא "לא", הצג דוגמה נגדית. אם התשובה היא "כן", הצג הוכחה (בד"כ ע"י אינדוקציה). זכור שהתכונה היא אסימפטוטית, ולכן הדוגמה הנגדית צריכה להכיל סדרה אינסופית של עצים במשפחה, ולא רק עץ יחיד.

א. כל צומת בעץ הוא או עלה, או שיש לו שני בנים.

פתרון: התכונה אינה מספיקה. ניתן להסתכל על עץ שרוך ימני, כשלכל צומת בשרוך הוספנו גם בן שמאלי. מספר הרמות הוא בערך $n/2$.

ב. הגודל של על תת עץ יכול להיכתב כ- $2^k - 1$, כאשר k הוא איזשהו מספר שלם (k אינו זהה לכל תת עץ).

פתרון: התשובה היא כן.

נתבונן בתת עץ כלשהו עם שורש r . אנחנו יודעים שגודלו של תת העץ הזה הוא $2^{k_1} - 1$, בנוסף גודלו של תת העץ הימני הוא $2^{k_2} - 1$, וגודלו של תת העץ השמאלי הוא $2^{k_3} - 1$. לכן, נקבל כי:

$$\begin{aligned}2^{k_1} - 1 &= 2^{k_2} - 1 + 2^{k_3} - 1 + 1 \\2^{k_1} &= 2^{k_2} + 2^{k_3}\end{aligned}$$

נראה שהנ"ל גורר כי $k_2 = k_3$. ברור כי $k_1 \geq k_2$ וגם $k_1 \geq k_3$. נניח בלי הגבלת כלליות ש- $k_2 \leq k_3$. נקבל אם כן: $2^{k_1 - k_2} = 1 + 2^{k_3 - k_2}$. הנ"ל גורר כי: $k_1 - k_2 = 1$, וגם $k_3 - k_2 = 0$. במילים אחרות, $k_2 = k_3$, ו- $k_1 = k_2 + 1$. נקבל אם כן כי כל קודקוד מכיל תת עץ ימני בגודל 2^k , ותת עץ שמאלי בגודל 2^k , והעץ מאוזן בצורה מושלמת.

ג. קיים קבוע $0 < c < 1$ כך ש, לכל צומת בעץ, גודל תת העץ הקטן יותר הוא לפחות c פעמים גודל תת העץ הגדול יותר.

פתרון: כן. ההוכחה היא באינדוקציה. נניח ששני תתי העצים של עץ x מגודל n הם y ו- z עם גדלים n_1 ו- n_2 בהתאמה. לפי הנחת האינדוקציה, הגובה של y הוא לכל היותר $d \log n_1$, והגובה של z הוא לכל היותר $d \log n_2$. נניח בלי הגבלת כלליות ש- $n_1 \geq n_2$. לכן, מתקיים כי $n_2 \geq cn_1$. לכן, אנו מקבלים: $n = n_1 + n_2 + 1 \geq (c + 1)n_1$. עומק העץ הוא:

$$d \log n_1 + 1 \leq d \log (n / (c + 1)) + 1 = d \log n - d \log (c + 1) + 1 \leq d \log n$$

כאשר $d \log (c + 1) \geq 1$. לכן, עבור d מספיק גדול, אי השיויון הנ"ל מתקיים, ונוכל לקבל את הדרוש.

ד. קיים קבוע c כך שלכל צומת של העץ, הגבהים של שני תתי העצים שונים בכלל היותר c .

פתרון: כן. זוהי למעשה הכללה של עצי AVL (שם ראינו כי ההבדל הוא 1).
 תהי $n(h)$ פונקציה המקבלת גובה של עץ, ומחזירה את המספר המינימלי של צמתים של עץ המקיים את התכונה שאנו מתבוננים בה. נראה באינדוקציה כי: $n(h) \geq (1 + \alpha)^h - 1$, עבור איזשהו קבוע $0 < \alpha \leq 1$. מכאן, נסיק כי: $h \leq \log_{1+\alpha}(n+1) = O(\log n)$.
 בסיס - תת עץ בגובה 0 יש צומת יחיד, ומתקיים: $1 \geq (1 + \alpha)^0 - 1$.
 צעד - נניח שהטענה נכונה עבור גובה h . כעת, נתבונן בעץ בגובה h , ונתבונן בשני תתי העצים שלו. אנחנו יודעים שגובהו של האחד הוא לכל היותר $h - 1$, ולכן של השני הוא לכל הפחות $h - 1 - c$. לכן, נקבל:

$$n(h) \geq n(h-1) + n(h-1-c) + 1$$

בעזרת טענת האינדוקציה, נקבל:

$$\begin{aligned} n(h) &\geq (1 + \alpha)^{h-1} - 1 + (1 + \alpha)^{h-1-c} - 1 + 1 \\ &\geq 2(1 + \alpha)^{h-1-c} - 1 = 2(1 + \alpha)^h \cdot (1 + \alpha)^{-(c+1)} - 1. \end{aligned}$$

כעת, נוכל לבחור α כך ש: $(1 + \alpha)^{c+1} < 2$, מה שאומר כי: $(1 + \alpha)^{-(c+1)} \geq 1/2$. נקבל אם כן:

$$n(h) \geq (1 + \alpha)^h - 1$$

ולכן צעד האינדוקציה מתקיים.

ה. הגובה הממוצע של כל צומת בעץ הוא $O(\log n)$.

פתרון: לא. נניח עץ שבו $n - \sqrt{n}$ צמתים הם ממש עץ מאוזן לחלוטין, ושאר \sqrt{n} הצמתים הם מעין שרוד. גובה העץ הוא: $\log(n - \sqrt{n}) + \sqrt{n} = \Omega(\sqrt{n})$, בעוד שממוצע גבהי הצמתים:

$$\begin{aligned} &(1/n) \cdot (\sqrt{n} \cdot (\sqrt{n} + \log n)) + (n - \sqrt{n}) \cdot \log n \\ &= (1/n) \cdot (n + n \log n) = O(\log n) \end{aligned}$$

תרגיל. עץ שלם הוא עץ שבו כל צומת הוא או עלה, או שיש לו שני בנים. הוכח באינדוקציה על מבנה העץ של עץ שלם בעל n עלים יש $n - 1$ קודקודים פנימיים.

הוכחה: בסיס: לעץ בעל עלה יחיד (1) אין צמתים פנימיים (0).

צעד: נניח שעבור כל העצים השלמים בעלי $k < n$ עלים יש $k - 1$ קודקודים פנימיים.

נוכיח עבור העצים השלמים בעלי n עלים:

יהי עץ שלם בעל n עלים. מכיוון שהוא עץ שלם, לכל קודקוד שני בנים, ובפרט לשורש. נתבונן בבנים של השורש. העץ הימני הוא עץ שלם בעל n_1 עלים, ו- $n_1 - 1$ קודקודים פנימיים. העץ השמאלי הוא עץ שלם בעל $n - n_1$ עלים, ובעל $n - n_1 - 1$ קודקודים פנימיים. לכן, בסה"כ הקודקודים הפנימיים בעץ (כולל השורש) הוא:

$$n_1 - 1 + n - n_1 - 1 + 1 = n - 1$$



והוכחנו את הטענה.

תרגיל. נתונים שני עצי חיפוש. כיצד ניתן למזג אותם לעץ חיפוש אחד?

פתרון: הרעיון הוא לבצע מעבר *infix*, ולקבל שני מערכים ממויינים. לאחר מכן, נבצע מיזוג של שני המערכים למערך אחד ממויין. לבסוף, נבנה עץ חיפוש מאוזן ממערך ממויין בצורה הבאה:

בניית עץ מאוזן למערך (A, i, j) . (בהתחלה, $j = n, i = 1$)

1. אם $i = j$, נחזיר עלה - $A[i]$. אחרת:

(א) נסמן $med = (i + j)/2$

(ב) ניקח את $A[med]$ להיות השורש r .

(ג) נבנה בצורה רקורסיבית עץ מאוזן עבור $(A, i, med - 1)$. נסמן את העץ המתקבל כ- T_1 .

(ד) נבנה בצורה רקורסיבית עץ מאוזן עבור $(A, med + 1, j)$. נסמן את העץ המתקבל כ- T_2 .

(ה) נבנה את העץ שבו r הוא השורש, הבן השמאלי הוא T_1 , הבן הימני הוא T_2 .

נקבל כי: $T(n) = 2T(n/2) + 1$, ולכן בסה"כ - $O(n)$.

תרגיל: הכנסה / הוצאה מ- *AVL* (דפים).

תרגיל: בנה מבנה נתונים, התומך בפעולות הבאות:

1. הכנסה.

2. הוצאה.

3. חיפוש.

4. מציאת האיבר ה- k בגודלו (כלומר, $k - 1$ מהאיברים קטנים ממנו, ושאר $n - k$ האיברים - גדולים ממנו).

נרצה מבנה נתונים התומך בכל הפעולות הנ"ל, בסיבוכיות $O(\log n)$ לכל פעולה.

פתרון: לוקחים *AVL* ומוסיפים לכל צומת שדה נוסף - "משקל" - כמה איברים נמצאים בתת העץ. בצורה כזו, אפשר