<u>שאלות על עצי חיפוש</u>

שאלה 1

הוכיחו שבכל עץ חיפוש בינארי המכיל n מפתחות שונים זה מזה, יש לפחות עלה אחד כך שאורך המסלול משורש העץ לעלה הוא $\Omega(\log(n))$

פתרון

נגדיר את אורך המסלול הארוך ביותר כ h ונשאל את השאלה הבאה: בעץ שגובהו (אורך המסלול הארוך ביותר משורש לעלה h, (כמה צמתים אפשר לשים לכל היותר?

נשים לב שאם יש צומת שאין לו שני ילדים, אפשר לבנות עץ גדול יותר על ידי הוספת ילד אליו. מכאן נקבל שהעץ המכיל את מספר הצמתים המקסימאלי מבין הצמתים שגובהם h, הוא עץ שבו **כל** מסלול משורש לעלה הוא בגובה h, ולכל צומת (למעט העלים) יש שני ילדים.

בעץ כזה, ברמה הראשונה יש 1 צומת, ברמה הבאה 2 צמתים, וברמה ה i יש 2^{i-1} צמתים. נקבל שמספר הצמתים הוא $\sum_{i=1}^{h} 2^{i-1} = 2^h - 1$.זה המספר ה**מקסימאלי** של הצמתים. לכן, עבור כל עץ שגובהו h ומספר צמתיו n בהכרח מתקיים $n \leq 2^h - 1 \leq 2^h$. מלקיחת $\log$ משני האגפים נקבל $h = \Omega(\log(n))$
.

שאלה 2

הוכיחו שכל אחת מהפעולות למעבר על כל איברי עץ שראינו (in order/pre order/ post order) פועלת בזמן ליניארי במספר האיברים בעץ.

פתרון

עלינו להראות שיש שני קבועים c_0, ו c_1 המקיימים $c_1 \cdot n \leq T(n) \leq c_0 \cdot n$ עבור ערכי n גדולים מספיק.

הטענה שקיים c_1 $T(n) \geq c_1 \cdot n$ כלומר $T(n) = \Omega(n)$ ברורה, מפני שאנו יודעים שהפונקציה עוברת על כל אחד מהצמתים, ויש n צמתים עפ"י ההגדרה. נוכיח לכן רק את הטענה לגבי קיומו של c_0.

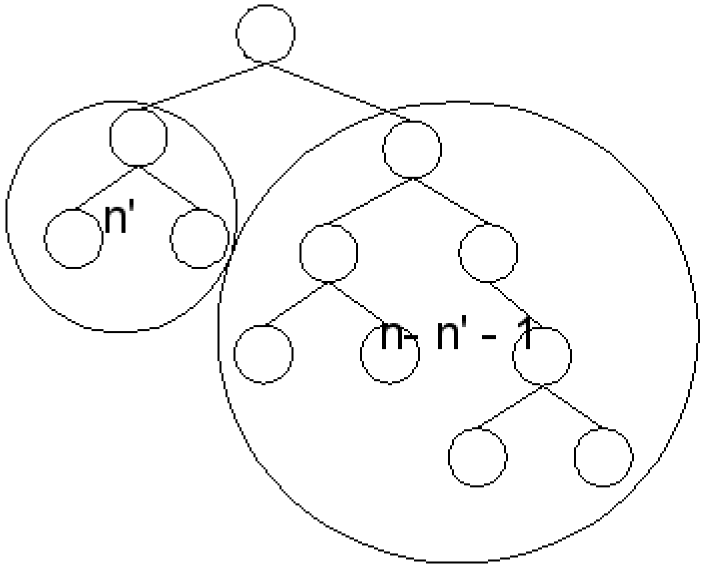
הוכחה : (בסיס האינדוקציה) נבחר את בסיס האינדוקציה כ $n = 1$.

ונסמן ב k את זמן הריצה של הפונקציה על קלט בגודל 1. מכאן נקבל את האילוץ על c_0: $k \leq c_0 \cdot 1 = c_0$. בוודאי שקיים c_0 המקיים אילוץ זה.

(מעבר האינדוקציה) נניח שהטענה מתקיימת לכל $n' < n$

נוסחת הנסיגה המתאימה לבעיה: $T(n) = T(n') + O(1) + T(n - n' - 1)$

קל לראות זאת גם מהתרשים הבא:



הנוסחה אומרת שקיים k' כך שמתקיים
$$T(n) \leq T(n') + k' + T(n - n' - 1)$$

נשתמש בנוסחת הנסיגה ובהנחת האינדוקציה, ונקבל

$$T(n) \leq T(n') + k' + T(n - n' - 1) \leq$$
$$c_0 \cdot n' + k' + c_0(n - n' - 1) =$$
$$c_0 \cdot n + (k' - c_0)$$

אם נבחר $k' \leq c_0$ אז $(k' - c_0)$ שלילי, ולכן בהכרח $T(n) \leq c_0 \cdot n$.

נשים לב שבסיס האינדוקציה הגביל אותנו לבחור $k \leq c_0$, ומעבר האינדוקציה הגביל אותנו לבחור $k' \leq c_0$.

מכאן יוצא שאכן יש c_0 המתאים לבעיה (למעשה יש אינסוף כאלה) - עלינו פשוט לבחור c_0 המקיים

$$c_0 \geq \max\{k, k'\}$$