נושאים מתקדמים באלג' - תרגיל ב

1 א. יש לתאר את אופן ביצוע פעולות הכנסה, שרשור ופיצול.

הכנסה: נתחיל בלחפש את המיקום הנכון של האיבר החדש x בעץ שלנו T: - נתבונן על שני מצביעים "האצבעות" בקצוות העץ. 10/10 (מלא)

- נתחיל במקביל על האצבעות הימנית והשמאלית לכוון השורש.
- באצבע הימנית נבדוק כל צומת ונבדוק האם x גדול מהמינימום הקטן ביותר בצומת. אם כן, נתחיל בחיפוש רגיל בעץ 2-4 כפי הלמדנו החל מצומת זה. אחרת נמשיך לעלות.
- באצבע השמאלית החיפוש סימטרי.
- אם הגענו לשורש משתי האצבעות, נתחיל בחיפוש רגיל מהשורש.

לאחר שמצאנו את המיקום הנכון של x, נכניס אותו כבן חדש. אם לאבא של x כבר היו 4 בנים, נבצע הפרדה כמו בעץ 2-4 רגיל, ובמקרה והצומת המופרד נמצא על אחת האצבעות, נתקן את האצבעות לשורש. נמשיך בתיקון העץ כלפי השורש תוך התחשבות באצבעות.

שרשור: נניח שכל איברי T_1 קטנים מאיברי T_2. נתחיל לטפס על האצבע הימנית של T_1 והשמאלית של T_2 במקביל, עד שנגיע לשורש של אחד העצים (הנמוך מביניהם). נניח בה"כ שהגענו לשורש של T_1 ראשון. נוסיף את T_1 (לאחר ניתוק את השורש של T_1) כבן של הצומת אליה הגענו בטיפוס על האצבע השמאלית של T_2 ונעשה תיקון כמו בהכנסה. בנוסף נעדכן את האצבע השמאלית של T_2 על ידי טיפול מהצומת אליה נוסף T_1 ומעבר על האצבע השנייה ומחיקת המצביעים, החל מצביע מהשורש הישן של T_1 לאבא החדש שלו, ועדכון תחילת האצבע השמאלית של העץ המאוחד להיות האצבע השמאלית של T_1, ומחיקת האצבע הימנית של T_1. צ"ע מתוין

פיצול: בהינתן x, נחפש אותו ב-T (כמפורט בהכנסה) ונזכור
את v - הצומת הראשון של אחת האצבעות שבה פנינו לראשונה
מהאצבע (למשל שמאלה של האצבע הימנית) במהלך החיפוש. כעת
נלך מ-x ל-v ונבנה שני עצים T_1,T_2: בכל צומת בדרך, ניקח
את הבנים שמשמאל למסלול ונוסיף אותם ל-T_1, ואת הבנים שמימין
למסלול נוסיף ל-T_2 (קל לראות שהתוספות נכונות ושהעצים המתקבלים
כשרשור בלי איברים גדולים אחד מהשני). לבסוף נבדוק מאיזה אצבע
הגיע v, נניח מהאצבע הימנית, נמחק את v מ-T ונתקן
את העץ כמו במחיקה בעץ 2-4, ונשרשר את T עם T_1
(הנחנו ש-v היה על האצבע הימנית) לאחר שנוסיף ל-T_1 ו-T_2
אצבעות. קיבלנו שני עצים כנדרש.

ב. 1) יש להראות ששרשור ופיצול עולים $O(\log n)$ במקרה הגרוע.

10/10

שרשור: החיפוש לוקח $O(\log n)$, הכנסה גם כן $O(\log n)$ (כמו בעץ
2-4 רגיל) ותיקון האצבעות ממיקום ההכנסה ועד לעלים גם
עולה $O(\log n)$, סה"כ $O(\log n)$.

פיצול: נראה כמה זמן נדרש לנו ליצירת T_1 (יצירת T_2 שקולה). נסתכל
על הבניה כעל תהליך איטרטיבי, מתחיל ב-x ומסתיים ב-v ובכל
איטרציה עולים רמה בעץ. באיטרציה i המרחק שלנו מ-x
הוא i, אנחנו נמצאים בצומת v_i, והגובה הנוכחי של T_1
הוא h_i (לאחר השרשור $\underbrace{\text{יהיה לכל}}_{\text{לכל היותר } i-1}$ היותר i). באיטרציה
זו מס' ההפרדות המירבי שנבצע הוא $(i-1)-h_i+1$
הגובה של הבנים של v_i · הגובה של T_1 · אחד נוסף אם השרשור
שזה $O(i-h_i)$. סה"כ במהלך הטיול
מ-x ל-v עלות הפיצולים היא $\sum_{i=0}^{h_v} i-h_i$
ששרשור באיטרציה i
ובסכום זה שווה לגובה של v, שיכול להיות
לכל היותר $O(\log n)$. לאחר שבנינו את T_1,T_2, להוסיף
להם אצבעות עולה $O(\log n)$, למחוק את v מ-T עולה $O(\log n)$

ולשרשר את T עם T_i עולה גם כן $O(\log n)$, סה"כ $O(\log n)$.

ב) יש להראות ששרשור לוקח $O(\log(\min\{n_1, n_2\}))$ כאשר משרשרים T_1 עם n_1 איברים עם T_2 עם n_2 איברים, ופיצול לוקח $O(\log(\min\{d_i, |T_i|-d_i\}))$ כאשר x_i האיבר ה-d_i הקטן ביותר ב-T_i.

שרשור: מציאת נק' השרשור עולה $O(\log(\min\{n_1, n_2\}))$ מכיוון שמתחילים מהעלים לאחד השורשים. תיקון הצמתים מתבצע מנק' השרשור ומעלה לכן גם עלותו זהה. נסמן ב-k את מס' ההפרדות שנאלצנו לעשות על מנת לתקן את מבנה העץ לאחר הכנסת השורש של העץ הקטן לעץ הגדול. נשתמש בפונ' הפוטנציאל שראינו בהרצאה $\Phi = 2\times\#_{4 \text{ בנים}}^{\text{צמתים עם}} + \#_{2 \text{ בנים}}^{\text{צמתים עם}}$.
ניזכר ש- $am(cat) = ac(cat) + \Delta\Phi$. ראינו כעת ש- $ac(cat) = O(\log \min\{n_1,n_2\}) + k$. כל הפרדה מורידה צומת עם 4 בנים אחד ומוסיפה צומת עם 2 בנים, לכן $\Delta\Phi = -2k+k = -k$ ומכאן ש-
$am(cat) = O(\log(\min\{n_1,n_2\})) + k - k = O(\log(\min\{n_1, n_2\}))$.

פיצול: הפיצול מתחיל במציאת x_i בעץ, מה שלוקח $O(\log(\min\{d_i, |T_i|-d_i\}))$.
בניית T_1 (וכן T_2) מורכבת מהרבה שרשורים. הדין שאם אנו יודעים את נק' השרשור (אין צורך לחפש אותה) אז עלותו $O(1)$ (כי עושים תיקון אחד כלא - מוסיפים צמתים ל- T_1, T_2 רק בסוף כשמגיעים ל-x). מציאת נק' השרשור במקרה זה עולה $h_i - i$ (לפי הסימונים מסעיף (1)) ולכן העלות של מציאת נק' השרשור בכל השרשורים במהלך בניית T_1 היא $O(\log(\min\{d_i, |T_i|-d_i\}))$ וזו גם העלות של בניית T_1 (וכן T_2). את נק' השרשור של T_1 עם T_i קל למצוא - המחליף של הורג של v לאחר מחיקת v. מחיקת v עולה $O(1)$ amortized לפי פונ' הפוטנציאל, לכן סה"כ העלות של הפיצול היא:

$$am(split) = find(x_i) + \sum_{T_1} am(cat) + \sum_{T_2} am(cat) + am(remove) + am(cat(T_i, T_1))$$
$$= O(\log(\min\{d_i, |T_i| - d_i\}))$$

ב. יש להראות שאם מתחילים עם עץ T עם n איברים ובכל פעם מפצלים את אחד העצים הקיימים לפי איבר x כלשהו, נקבל זמן ריצה $O(n)$. 10/10

נניח שאנו מפצלים עד שנגיע ל-n עצים בגודל 1 - המקרה הגרוע ביותר. אנו יודעים ש- $am(split) = ac(split) + \Delta\Phi$. נסכום על כל פעולות הפיצול: $\sum am(split) = \sum ac(split) + \Phi_n - \Phi_0$, ונחלץ את העלות הכוללת: $\sum ac(split) = \sum am(split) + \Phi_0 - \Phi_n$. מן הסעיף הקודם אנו יודעים ש- $am(split(T',x)) = O(\log \min\{d_i, |T'|-d_i\})$ כאשר d_i המרחק של x מהאיבר השמאלי ב-T'. המקרה הגרוע ביותר הוא אם x באמצע, ואז מחיר הפיצול $O\left(\log \frac{|T'|}{2}\right)$:

$$\sum am(split) \le \sum_{i=1}^{\log n} 2^{i-1} \log\left(\frac{n}{2^i}\right) = \sum_{i=1}^{\log n} 2^{i-1}(\log n - i) \ominus$$

נסמן $\log n - i = j$ $\quad (i = \log n - j)$

$$\ominus \sum_{i=1}^{\log n} 2^{\log n - j - 1} \cdot j = \frac{n}{2} \sum_{i=1}^{\log n} 2^{-j} \cdot j = \frac{n}{2} \sum_{j=\log n - 1}^{0} 2^{-j} \cdot j =$$

(אין חשיבות לסדר סכימה) $= \frac{n}{2} \sum_{j=0}^{\log n - 1} 2^{-j} \cdot j \underset{\uparrow}{=} O(n)$

$$\sum_{i=0}^{\infty} i 2^{-i} = O(1)$$

מכאן ש-

$$\sum ac(split) = \sum am(split) + \Phi_0 - \Phi_n \le O(n) + \Phi_0 - \Phi_n = O(n)$$

בהתחלה במקרה הגרוע ביותר כל הצמתים הן עם 4 בנים, לכן $\Phi_0 = 2n$

בסוף כל העצים בגודל 1 לכן $\Phi_n = 0$

2 א. 10/10

מוגדר $w(i,j) = \sum_{l=i}^{j} p_l$, $B[i,i] = 0 \quad i = 1,\dots,n$

$$B[i,j] = w(i,j) + \min_{i<t\le j}\{B_{i,t-1} + B_{t,j}\}$$

יש להראות ש- $B[1,n]$ נותן את עלות העץ האופטימלי, ולמצוא אלג' לחישוב זה בזמן $O(n^3)$.

אינטואיטיבית קל לראות ש- $w(1,n)$ זו העלות המינימלית להכנסת n האיברים וזאת שכן $d_i \ge 1$ לכל i. אחר כך אנחנו עוברים על כל האפשרויות לפצל בין שני הבנים את כל האיברים תוך התחשבות בתשלום של כל פיצול ובוחרים את האפשרות הזולה ביותר. פורמלית נוכיח באינדוקציה.

נתחיל באלכסון, עבור $B[i,i]=0$ העץ הוא צומת יחיד ועבור אופטימלי.

עבור $j=i+1$ האפשרות היחידה היא  ולכן זה האופטימלי

וזה אכן העץ של $B[i,j]= \underbrace{W(i,j)}_{=p_i+q_j}+\underbrace{B[i,i]}_{=0}+\underbrace{B[j,j]}_{=0}$

עתה נניח עבור $B[i,i],\dots,B[i,j-1]$ ועבור $B[j,j],\dots,B[i+1,j]$ ונוכיח עבור $B[i,j]$. (ההנחה היא ש $B[k,l]$ נותן את העץ האופטימלי עבור $(k,\dots,l)$. אנחנו עוברים על כל אפשרויות החלוקה כאשר לכל אפשרות חלוקה אנו יודעים לפי הנחת האינדוקציה את העלות האופטימלית לפיה ומוסיפים $W(i,j)$ - שזה הוספת שורש משותף לשני תתי העצים ולכן המינימום מבין האפשרויות יתן לנו את החלוקה והעץ האופטימליים.

על מנת לחשב את $B[1,n]$, נתחיל מהאלכסון, $B[i,i]$ ונמשיך לחשב את האלכסונים שמעל, עד שנגיע ל- $B[1,n]$. בכל פעם שנרצה לחשב את $B[i,j]$ פשוט נעבור על $B[i,i]+B[i+1,j],\dots,B[i,j-1]+B[j,j]$ כי כל האיברים הללו נמצאים על אלכסונים שכבר חישבנו ונמצא את המינימום, מה שעולה לנו $O(n)$. יש $O(n^2)$ איברים מעל האלכסון, לכן החישוב יעלה $O(n^3)$. על מנת לבנות את העץ פשוט לכל $B[i,j]$ נזכור את ה-t שהביא את הסכום למינימום ונפצל את הצמתים לפי t זה. נתחיל מ $B[1,n]$ ונבנה את העץ לבנין הילדים לפי ה-t-ים ששמרנו.

ב. יש להראות של $B[i,j]$ יש את תכונת המונג': $B[i,j]+B[i',j'] \le B[i',j]+B[i,j']$.

נוכיח באינדוקציה. אנו יודעים ש $1\le i\le i'\le j\le j'\le n$.

עבור $i=j'$ נקבל את אותו האיבר 4 פעמים באי-שוויון - טריוויאלי.

עבור $j'-i=1$ יש שני מקרים:

I. $i=i'$, $j=j'$ כל האיברים אותם, טריוויאלי.

II. $i=j$ או $i'=j'$ אז נקבל ששני איברים מתאפסים (אחד מכל צד) והשניים

10/10

האחרים זה אותו אידר.

נניח נכונות עבור $k > i'-j$ ונוכיח עבור $k = i-j'$.

נתבונן על שני האידרים שסכומם צריך להיות גדול שווה לשני האחרים, ונסמן:

$$t_1 = \underset{i < t \le j'}{\arg\min}\{B[i, t-1] + B[t, j']\}, \quad t_2 = \underset{i' < t < j}{\arg\min}\{B[i', t-1] + B[t, j]\}$$

נניח ש $t_1 \le j$ (בה"כ אחרת במקרה השני). מההגדרה:

$$B[i,j] \le w(i,j) + B[i,t_1-1] + B[t_1,j]$$

$$B[i',j'] \le w(i',j') + B[i',t_2-1] + B[t_2,j']$$

נסכום את המשוואות ונקבל:

$$B[i,j] + B[i',j'] \le w(i,j) + w(i',j') + B[i,t_1-1] + B[t_1,j] + B[i',t_2-1] + B[t_2,j'] \Leftrightarrow$$

$$\left(w(i,j) + w(i',j') = \sum_{l=i}^{j} p_l + \sum_{l=i'}^{j'} p_l = \sum_{l=i}^{j'} p_l + \sum_{l=i'}^{j} p_l = w(i,j') + w(i',j) \right)$$

$$\left(\text{(שינוי סדר סכימה)} \quad i \underbrace{\;\times 1\; i' \;\times 2\; j \;\times 1\;}\, j' \right)$$

$$\Leftrightarrow \underbrace{w(i,j') + B[i,t_1-1]}_{= B[i,j'] - B[t_1,j']} + \underbrace{w(i',j) + B[i',t_2-1]}_{= B[i',j] - B[t_2,j]} + B[t_1,j] + B[t_2,j'] =$$

$$= B[i,j'] - B[t_1,j'] + B[i',j] - B[t_2,j] + B[t_1,j] + B[t_2,j'] \le B[i,j'] + B[i',j]$$

צריך בשביל המעבר הזה להראות ש $B[t_1,j] + B[t_2,j'] - B[t_1,j'] - B[t_2,j] \le 0$

וזה נכון מהנחת האינדוקציה.

עבור $t_1 > j$ נשתמש ב-

$$\begin{cases} B[i',j'] \le w(i',j') + B[i',t_1-1] + B[t_1,j'] \\ B[i,j] \le w(i,j) + B[i,t_2-1] + B[t_2,j] \end{cases}$$

.

ב. עבור $1 \le d \le n$, $1 \le i < t \le i+d \le n$ נגדיר את המטריצה B^d:

$$B^d[i,t] = w(i,i+d) + B[i,t-1] + B[t,i+d]$$

יש להראות ש B^d מקיימת את תכונת מונג בחלק המוגדר ולהראות איך

למצוא מינימום בשורות בעזרת SMAWK.

נתבונן באיברים עבור $1 \le i \le i' < j \le j' \le i+d \le n$: (הפולסמן אינו מוגדר)

$$B^d[i,j] + B^d[i',j'] - B^d[i',j] - B^d[i,j'] = w(\cancel{i,i+d}) + w(\cancel{i',i'+d}) - w(\cancel{i',i'+d}) - w(\cancel{i,i+d}) +$$

10/10

$$+ B[i, j-1] + B[j, i+d] + B[i', j'-1] + B[j', i'+d] - B[i', j-1] - B[j, i'+d]$$
$$- B[i, j'-1] - B[j', i+d] \leq 0$$

סכום המחוברים המוקפים ≥ 0 מתכונת המונג' על B

ועוד סכום המחוברים הלא מוקפים.

לכן תכונת המונג' מתקיימת.

כעת על מנת לחשב את המינימום בכל שורה, נכפיל פי (-1) את ערכי המטריצה המוגדרים, ונשלים את החלקים שמעל ומתחת לחלק המוגדר: מעליו נשים בכל המקומות $-2n-2$ מכיוון ש $\Sigma d_{i,j} \leq n$ וכל איבר מוגדר הוא לכל היותר סכום שני "צדדים" של עצים $+1$, ומתחתיו נשים $-2n-2$ באלכסון הראשון מתחת לאלכסון הראשי, $-2n-3$ באלכסון שאחריו וכן הלאה. כעת כל המטריצה מקיימת את תכונת המונג', אפשר להריץ SMAWK והוא יתן לנו את המקס' בכל שורה, שהוא המינימום לפני ההכפלה פי (-1).

ב. יש להראות איך בעזרת סעיף ג' נמצא את העץ האופטימלי ב $O(n^2)$ זמן.   10/10

נעבור באיטרציות עבור $d=1,\dots,n-1$. בכל איטרציה נחשב SMAWK של B^d ונשמור את המינימום של השורות שמצאנו. עבור $d=1$, נקבל את איברי B^d ע"י חישוב B מכיוון שזה קל. עבור $d=2,\dots,n-1$ כשנרצה לדעת איבר כלשהו של B^d (לא נחשב את כל המטריצה אלא רק איברים לפי בקשת ה-SMAWK) נשתמש במינימומים מהחישובים הקודמים כדי לדעת את ערכי B המרכיבים את האיבר של B^d. סה"כ נריץ $O(n)$ SMAWK-ים שעלות כל $O(n)$, סה"כ $O(n^2)$. על מנת לבנות את העץ נזכור מאיזו עמודה ושורה הגיע כל מינימום ולכך נוכל להרכיב את העץ, וזה יעלה לנו $O(n)$.

זמן ריצה כולל $O(n^2)$.

③ א. יש להראות כיצד עונים על שאילתא בזמן $O(\log n + k\log\log n)$.

$\frac{9}{10}$

בהינתן $x_1 \leq x_2$, נוסיף את כל תתי העצים בזמן החיפוש של

x_1 ו- x_2 בעץ הראשי (נוסיף אחים ימניים בכל פניה שמאלה

בחיפוש של x_1 ואחים שמאליים בכל פניה ימינה בחיפוש של x_2)

ונקבל $O(\log n)$ עצי finger עם איברים צריכים. ניקח

מכל עץ את האיבר הגדול ביותר בזמן $O(1)$ לפי המצביע

הימני ונבנה מהם עץ מאוזן (לא ממוין) בזמן $O(\log n)$ באופן

הבא: נחלק את האיברים לזוגות, לכל זוג נוסיף אבא

משותף ונצבור באבא איזה איבר מהבנים שלו גדול יותר. נמשיך

כך על הרמות שיצרנו עד שנישאר עם שורש יחיד. הזמן שנדרש

לבנות את העץ הזה $\log n + \frac{1}{2}\log n + \frac{1}{4}\log n + \dots + \frac{1}{2^{\log\log n}}\log n = O(\log n)$. אפשר גם להשתמש בערימה

כעת נעשה k extract-max מהעץ שבנינו, כאשר מציאת

המקסימום לוקחת $O(\log\log n)$ - כעומק העץ, ולאחר הוצאת האיבר

המירבי נלך לעץ שממנו הוא הגיע, ניקח את האיבר הבא הגדול, (בזמן $O(1)$) כיצד נעשה זאת בזמן $O(1)$?

נכניס במקום האיבר שהוצאנו (לאיבר הגדול) ונתקן את ההדבקות עד לשורש

בזמן $O(\log\log n)$. סה"כ זמן הריצה $O(\log n + k\log\log n)$

בניית עץ המקסימומים   k פעמים extract-max

ג. יש להראות שלכל צומת פנימי יש לכל היותר 32 ולפחות 2 בנים. 10/10

נניח שיש צומת פנימי (בגובה h) עם בן אחד בלבד. גודל התת עץ של

הבן שלו הוא לכל היותר $2\cdot 8^{h-1}$ שזה $\frac{8^h}{4}$ בסתירה לכך

שגודל הצומת עם הבן היחיד צריך להיות $\geq \frac{8^h}{2}$.

נניח שיש צומת פנימי בגובה h עם יותר מ-32 בנים. גודל

כל תת עץ של כל אחד מהבנים $\geq \frac{8^h}{16} = \frac{8^{h-1}}{2}$, מכאן שגודל

תת העץ של הצומת בסך הילדים $> 32\cdot\frac{8^h}{16} = 2\cdot 8^h$ בסתירה

לדרישה (2).

ד. יש לתאר הכנסה והוצאות לעץ כפי שתואר ב-ג'. יש לתכנן אותן כך שסדרה 10/10

של m פעולות תעלה $O(m\log^2 n)$.

הכנסה: בכל צומת נשמור את גודל תת העץ שלו. כשנרצה להכניס

איבר, נמצא את מיקומו בעץ ונכניס אותו כמו בעץ 2-4 רגיל,

ואז נעלה עד השורש תוך עדכון הגדלים של תתי-העצים השמורים

בכל צומת במסלול. כאשר הגודל של אחד הצמתים יהיה יותר

מ $2\cdot 8^h$ (את h נדע כי אנחנו מתחילים מלמטה ועולים לכוון השורש) נבצע פיצול

של הצומת כמו בעץ 2-4 כך שהחלוקה תהיה שווה ככל שאפשר. נמשיך

בתיקון עד לשורש. עדכון עצי האיזון המשניים זהה למקרה שבו

העץ הראשי היה עץ 2-4.

מחיקה: נחפש את האיבר שיש למחוק כמו בעץ 2-4 רגיל ונמחק אותו.

לאחר המחיקה יש לעלות לכוון השורש, לעדכן את המונים השמורים

עם גדלי תתי העצים, ובמקרה שבצומת כלשהו המונה ירד מ $\frac{8^h}{2}$, אחד

את הצומת עם אח ימני או שמאלי שלו (הקטן מביניהם) והמשך לבצע

את התיקון הפל מהצומת המאוחד (ייתכן שנצטרך לפצל אותו כדי לקבל

שני צמתים מאוזנים יותר) לכוון השורש. תחזוקת העצים המשניים

זהה, למקרה שבו העץ הראשי היה עץ 2-4.

על מנת להראות את העלות של רצף של m פעולות עדכון, נגדיר

פונקציית פוטנציאל: $\Phi = 2\cdot\sum_{v \mid |v|\le 8^h} \left||v|-8^h\right| + \sum_{v \mid |v|>8^h} \left||v|-8^h\right|$.

כמו שניתן לראות מה העלות של פעולת פיצול ואיחוי amortized. העלות

של הפיצול או האיחוי של הצמתים המשני' היא n. השינוי בפונ'

הפוטנציאל בפיצול או באיחוי הוא $-n$ גם כן, לכן העלות

פיצול ואיחוי

של הפעולות על הצמתים המשניים מתקזזת. נותרנו עם העלות של

ההכנסה/הוצאה שהיא $O(\log n)$. בכל הכנסה/הוצאה של נק'

עלינו לטייל ב-$O(\log n)$ צמתים משניים עד שנגיע לשורש ובכל

אחד לבצע עדכון שעלותו $O(\log n)$, סה"כ $O(\log^2 n)$, ובסה"כ של

m פעולות עדכון נשלם $O(m\cdot\log^2 n)$.