29.11.16

אלגוריתמים - הרצאה 2

מיון מהיר: $QS(A,p,r)$

נסתכל על פסאודו קוד של מיון מהיר:

```
QS(A,p,r)            T(n)
  if r>p
     q ← partition(A,p,r)
     QS(A,p,q-1)  → T(q-1)
     QS(A,q+1,r)  → T(n-q)
```

עצם מספר ההשוואות שמבצעים ב- partition : $n+1$

במקרה הגרוע המערך ממוין ונקבל $O(n^2)$

המקרה הטוב שבכל פעם נקרא לאיבר החציוני.

במקרה הממוצע האלגוריתם אקראי.

קלט: תמורות באורך n

כדי למצוא את זמן הריצה של המקרה הממוצע נמצא את התוחלת:

X- משתנה מקרי כאשר דוגמ שלנו $X=$ מס' ההשוואות שמיון מהיר מבצע על קלט באורך n

$X_q=$ מס' ההשוואות שמיון מהיר מבצע במקרה שה- pivot מגיע למקום ה- q.

$P_q=\frac{1}{n}=$ ההסתברות ש- X מגיע ל- X_q.

ולכן התוחלת: $E(x)=\sum X_q\cdot P_q$

$T(n)=$ מס' ההשוואות שמיון מהיר מבצע במקרה הממוצע.

$$T(n)=n+1+\sum_{q=1}^{n}\frac{1}{n}\left(T(q-1)+T(n-q)\right)=n+1+\frac{1}{n}\sum_{q=1}^{n}\left(T(q-1)+T(n-q)\right)=n+1+\frac{2}{n}\sum_{q=1}^{n}T(q-1)$$

$$\boxed{\begin{array}{c} T(0)+T(n-1)\\ T(1)+T(n-2)\\ \vdots\\ T(n-1)+T(0)\end{array}}$$

$$T(n)=n+1+\frac{2}{n}\sum_{q=1}^{n}T(q-1)\quad /\cdot n$$

(A) $n\cdot T(n)=n(n+1)+2\sum_{q=1}^{n}T(q-1)$ / הצבה עבור $n-1$:

(B) $(n-1)T(n-1)=(n-1)n+2\sum_{q=1}^{n-1}T(q-1)$

(A)-(B) $nT(n)-(n-1)T(n-1)=2n+2T(n-1)$

$nT(n)=2n+(n+1)T(n-1)$ $/_{n(n+1)}$ נחלק ב-

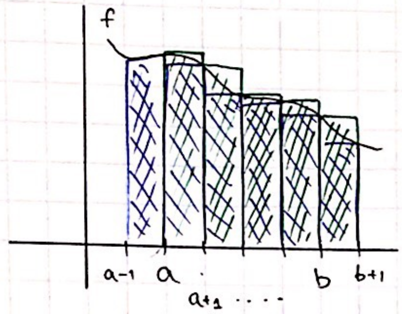
$$\frac{T(n)}{n+1}=\frac{2}{n+1}+\frac{T(n-1)}{n}$$

נציב: $G(n)=\frac{T(n)}{n+1}$ $\quad (G(0)=G(1)=0)$

$$G(n)=\frac{2}{n+1}+\frac{T(n-1)}{n}=\frac{2}{n+1}+\frac{2}{n}+G(n-2)=\frac{2}{n+1}+\frac{2}{n}+\frac{2}{n-1}+G(n-3)=\frac{2}{n+1}+\frac{2}{n}+\frac{2}{n-1}+\dots+\frac{2}{3}+\overbrace{G(1)}^{0}=2\sum_{i=3}^{n+1}\frac{1}{i}$$

בנוסף.

מהו הסכום $2\sum_{i=3}^{n+1}\frac{1}{i}$



$$\int_{a}^{b+1} f(x)\,dx \geq \sum_{i=a}^{b} f(i) \geq \int_{a-1}^{b} f(x)\,dx$$

לכן,

$$\ln(2+3)-\ln(3) = \int_{3}^{n+2}\frac{di}{i} \geq \sum_{i=3}^{n+1}\frac{1}{i} \geq \int_{2}^{n+1}\frac{di}{i} = \ln(n+1)-\ln(2)$$

$$T(n)=(n+1)G(n)=\Theta(n\log(n)) \Longleftarrow G(n)=C\cdot\log(n) \Longleftarrow$$

חסמים תחתונים למיונים <u>מבוססי השוואות</u>:

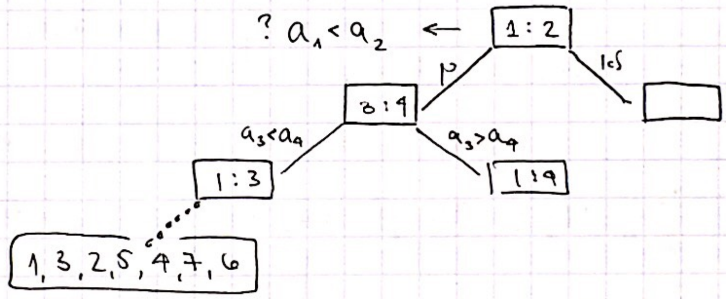
קלט: תמורה באורך n. (יש $n!$ קלטים אפשריים)

עבור אלגוריתם מיון מבוסס השוואות A, ניתן להגדיר עץ החלטות = עץ בינארי, שבו

לכל ק' יש 2 בנים בהכרח!

שמוגדר לפי ההשוואות ש-A מבצע.

T_A:



לכל אלגוריתם A מותאם עץ החלטות T_A יחיד מכיוון שבכל השוואה בהכרח הוא ילך לכיוון מסויים.

מס ההשוואות ש-A מבצע על קלט באורך n <u>במקרה הגרוע</u> =

עומק של $T_A \geq \log(n!) = \Theta(n\log(n))$

עומק עץ מאוזן הוא log של מס העלים.
בעץ כאן יש $n!$ קלטים אפשריים ולכן $n!$ עלים.

במקרה הממוצע נצטרך להסתכל על תוחלת מס ההשוואות ש-A מבצע.

$$E(X)=\sum x_i p_i = \sum_{i=1}^{n!}\left(\begin{matrix}\text{מס ההשוואות}\\ \text{ש-}A\text{ מבצע על}\\ \text{תמורה } x_i\end{matrix}\cdot p_i\right) = \frac{1}{n!}\cdot\sum_{i=1}^{n!}\left(\begin{matrix}\text{עומק העלה}\\ \text{ב-}T_A \quad x_i\end{matrix}\right)$$

$p_i = \frac{1}{n!}$ הנחנו התפלגות אחידה.

נסמן: $D(n)=\sum_{x \text{ עלים ב-}T}\left(\begin{matrix}\text{עומק העלה}\\ x \text{ ב-}T\end{matrix}\right)$.

נגדיר

$$d(m)=\min\{D(T)\}$$

T עץ עם m עלים.

24.11.16 בס"ד

דוגמאות:

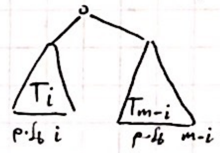
$$d(1)=\min\{D(1)\}=0$$

$$d(2)=\min\{D(\ )\}=2$$

$$d(3)=\min\{D(\ )\}=5$$

$$d(4)=\min\{\underbrace{D(\ )}_{8},\underbrace{D(\ )}_{9}\}=8$$

:T



$$D(T)=D(T_i)+D(T_{m-i})+m \Leftarrow$$

טענה: $d(m)=\min\limits_{1\le i\le m-1}\{d(i)+d(m-i)+m\}$

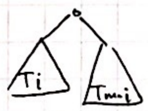
הוכחה -

$(\le)$: נבחר i כלשהו: לפי ההגדרה קיים עץ T_i (עם i עלים) כך ש- $D(T_i)=d(i)$.

כמו כן, קיים עץ T_{m-i} (עם $m-i$ עלים) כך ש- $D(T_{m-i})=d(m-i)$.

נבנה עץ T^* בצורה הבאה:

T^*



$$d(m)\le D(T^*)=D(T_i)+D(T_{m-i})+m\underset{\text{בחירה שלנו.}}{=}d(i)+d(m-i)+m$$

כיוון שזה נכון לכל i $\Leftarrow$ $d(m)\le\min\limits_{1\le i\le m-1}\{d(i)+d(m-i)+m\}$

$(\ge)$: קיים עץ T^* עם m עלים כך ש- $D(T^*)=d(m)$ ב-T^* נסמן את תת העץ השמאלי.

ב- T_j (עם j עלים) ואת תת העץ הימני ב- T_{m-j} (עם $m-j$ עלים)

$$d(m)=D(T^*)=D(T_j)+D(T_{m-j})+m\ge d(j)+d(m-j)+m\ge\min\limits_{1\le i\le m-1}\{d(\overset{i}{\cancel{j}})+d(m-\overset{i}{\cancel{j}})+m\}$$

$\blacksquare$

טענה: $d(m)\ge m\log(m)$

הוכחה - באינדוקציה

בסיס: $m=1,2,3,4$

נניח נכונות לטענה לכל $<m$ ונוכיח ל- m.

$$d(m)=\min\{d(i)+d(m-i)+m\}\underset{\text{הנחת האינדוקציה}}{\ge}\min\{i\log i+(m-i)\log(m-i)+m\}$$

$$f=i\log(i)+m\log(m-i)-i\log(m-i)+m$$

$$\frac{df}{di}=\log i+i\cdot\frac{1}{i}-\frac{m}{m-i}-\left(\log(m-i)-\frac{i}{m-i}\right)=\log i+1+\frac{-(m-i)}{m-i}-\log(m-i)=\log(i)-\log(m-i)$$

סוף.

הנגזרת שווה ל-0 $\Leftarrow$ $\log(i)=\log(m-i) \Leftrightarrow i=m-i \Leftrightarrow i=\frac{m}{2}$

$$\Leftarrow \quad \min_{1\le i\le m-1}\{i\log i+(m-i)\log(m-i)+m\}=\frac{m}{2}\log\frac{m}{2}+\frac{m}{2}\log\frac{m}{2}+m=m\log\left(\frac{m}{2}\right)+m$$

$$=m(\log(m-1))+m=m\log(m)$$

$\square$

לכל אלגוריתם A

$$E(A\ \text{מס' השוואות})=\frac{1}{n!}D(T_A)\ge\frac{1}{n!}d(n!)\ge\frac{1}{n!}n!(\log(n!))=\log(n!)=\Theta(n\log n)$$