

Beweis von c):

Wir bezeichnen im Folgenden mit H_{2h}^- den Baum, der entsteht, wenn wir H_{2h} so zerlegen, dass alle Elemente oberhalb des Zentrums wegfallen. Mit H_{2h+1}^+ bezeichnen wir den Baum, der entsteht, wenn wir H_{2h+1} so zerlegen, dass alle Elemente unterhalb des Zentrums wegfallen.

Lemma 86

H_{2h}^- und H_{2h+1}^+ haben jeweils 2^h Knoten. Bei der Herstellung aus H_{2h} bzw. H_{2h+1} werden $2^h - 1$ bzw. $2^{h+1} - 1$ Kanten aufgebrochen. Die wegfallenden Teile haben die Form H_s , $s < 2h$ bzw. $s < 2h + 1$.

Beweis:

Durch Induktion über r .

Induktionsanfang: für H_0 und H_1 gilt die Behauptung.

Induktionsannahme: die Behauptung gilt für alle H_p , $p < r$.

- 1 Sei $r = 2h$, $h > 0$. Wir betrachten die Partitionierung von H_{2h} mit Zentrum z wie in Lemma 85. Die Unterbäume $H_1, H_3, \dots, H_{2h-1}$ haben ihre Zentren oberhalb von z . Wir trennen sie von H_{2h} , indem wir h Kanten aufbrechen. Die abgetrennten Teile haben offensichtlich die Form H_s , $s < 2h$. Bei den Unterbäumen $H_0, H_2, \dots, H_{2h-2}$, mit Zentren unterhalb von z , wenden wir jeweils die Induktionsannahme an, d.h. wir erzeugen $H_0^-, H_2^-, \dots, H_{2h-2}^-$. Als Ergebnis erhalten wir H_{2h}^- .

Beweis (Forts.):

Damit gilt für die Zahl der aufzubrechenden Kanten $K^-(2h)$ zur Herstellung von H_{2h}^- :

$$K^-(2h) = h + \sum_{i=0}^{h-1} K^-(2i) \stackrel{I.A.}{=} h + \sum_{i=0}^{h-1} (2^i - 1) = \sum_{i=0}^{h-1} 2^i = 2^h - 1.$$

Für die Zahl $E^-(2h)$ der Elemente in H_{2h}^- gilt:

$$E^-(2h) = 1 + \sum_{i=0}^{h-1} E^-(2i) \stackrel{I.A.}{=} 1 + \sum_{i=0}^{h-1} 2^i = 1 + \underbrace{\sum_{i=1}^h 2^{i-1}}_{2^h - 1} = 2^h.$$

Beweis (Forts.):

- ② Sei $r = 2h + 1$, $h > 0$. Wir betrachten die Partitionierung von H_{2h+1} mit Zentrum z wie in Lemma 85. Die Unterbäume H_0, H_2, \dots, H_{2h} haben ihre Zentren unterhalb von z . Wir trennen sie von H_{2h+1} , indem wir $h + 1$ Kanten aufbrechen. Die abgetrennten Teile haben offensichtlich die Form H_s , $s < 2h + 1$. Bei den Unterbäumen $H_1, H_3, \dots, H_{2h-1}$, mit Zentren oberhalb von z , wenden wir jeweils die Induktionsannahme an, d.h. wir erzeugen $H_1^+, H_3^+, \dots, H_{2h-1}^+$. Als Ergebnis erhalten wir H_{2h+1}^+ . Damit gilt für die Zahl der aufzubrechenden Kanten $K^+(2h + 1)$ zur Herstellung von H_{2h+1}^+ :

Beweis (Forts.):

$$\begin{aligned}K^+(2h+1) &= h+1 + \sum_{i=1}^h K^+(2(i-1)+1) \\ &\stackrel{I.A.}{=} h+1 + \sum_{i=1}^h (2^i - 1) = 1 + \sum_{i=1}^h 2^i \\ &= 1 + \underbrace{\sum_{i=1}^{h+1} 2^{i-1}}_{2^{h+1}-1} - 1 = 2^{h+1} - 1.\end{aligned}$$

Für die Zahl $E^+(2h+1)$ der Elemente in H_{2h+1}^+ gilt:

$$E^+(2h+1) = 1 + \sum_{i=1}^h E^+(2(i-1)+1) \stackrel{I.A.}{=} 1 + \underbrace{\sum_{i=1}^h 2^{i-1}}_{2^h-1} = 2^h.$$

□

Beweis von d):

Lemma 87

Falls $k \leq 2^h - 1$, dann kann H_{2h} so zerlegt werden, dass die Komponente des Zentrums genau $2k + 1$ Elemente enthält, k davon über und k unter dem Zentrum. Dazu müssen $\leq 3k + 2h$ Kanten entfernt werden. Die entfernten Teile sind von der Form H_s , $s < 2h$.

Beweis:

Betrachte die Binärdarstellung von $k = k_0 2^0 + k_1 2^1 + \dots + k_{h-1} 2^{h-1}$ und die Partitionierung von H_{2h} mit Zentrum z wie in Lemma 85.

Beweis (Forts.):

Für jedes i mit $k_i = 1$, betrachte $H_{2^{i+1}}$ aus der Sequenz $H_1, H_3, \dots, H_{2^{h-1}}$ von Unterbäumen, deren Zentren oberhalb von z liegen, und schneide alle Elemente aus $H_{2^{i+1}}$, die kleiner als sein Zentrum sind (bilde also $H_{2^{i+1}}^+$). Dazu müssen höchstens $2k$ Kanten aufgebrochen werden, denn jedes $k_i = 1$ steht für 2^i in k , kostet aber nach Lemma 86 $K^+(2i+1) = 2^{i+1} - 1$ Kanten, also:

$$\sum_{i=0}^{h-1} k_i K^+(2i+1) \leq 2k.$$

Für jedes i mit $k_i = 0$, schneide $H_{2^{i+1}}$ ganz weg. Dabei werden $\leq h$ Kanten aufgebrochen. Genau k Elemente oberhalb z bleiben zurück, da jedes $k_i = 1$ für 2^i in k steht, und ein $H_{2^{i+1}}^+$ genau $E^+(2i+1) = 2^i$ Elemente enthält, also:

$$\sum_{i=0}^{h-1} k_i E^+(2i+1) = k.$$

Beweis (Forts.):

Für jedes i mit $k_i = 1$, betrachte H_{2^i} aus der Sequenz $H_0, H_2, \dots, H_{2^{h-2}}$ von Unterbäumen, deren Zentren unterhalb von z liegen, und schneide alle Elemente aus H_{2^i} , die größer als sein Zentrum sind (bilde also $H_{2^i}^-$). Dazu müssen höchstens $k - 1$ Kanten aufgebrochen werden, denn jedes $k_i = 1$ steht für 2^i in k und kostet uns nach Lemma 86 $K^-(2^i) = 2^i - 1$ Kanten, also:

$$\sum_{i=0}^{h-1} k_i(2^i - 1) \leq k - 1.$$

Für jedes i mit $k_i = 0$, schneide H_{2^i} ganz weg. Dabei werden höchstens h Kanten aufgebrochen. Genau k Elemente unterhalb von z bleiben zurück, da jedes $k_i = 1$ für 2^i in k steht, und ein $H_{2^i}^-$ genau $E^-(2^i) = 2^i$ Elemente enthält, also:

$$\sum_{i=0}^{h-1} k_i E^-(2^i) = k.$$

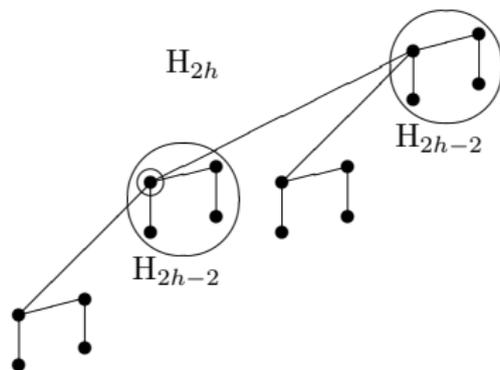
Beweis (Forts.):

Damit ergibt sich für die Gesamtanzahl aufzubrechender Kanten eine obere Schranke von $3k + 2h$. Lemma 86 liefert uns darüber hinaus die gewünschte Aussage über die Form der abgetrennten Teile. □

Beweis von d):

Betrachte H_{2h} .

- „größer“: $H_{2h-1}, H_{2h-3}, \dots, H_1$
- „kleiner“: $H_{2h-2}, H_{2h-4}, \dots, H_0$



$U(h) :=$ Anzahl der Elemente in $H_{2h} \geq$ Zentrum:

$$U(h) = 2U(h-1) = 2^h; \quad U(0) = 1$$

$D(h) :=$ Anzahl der Elemente in $H_{2h} \leq$ Zentrum:

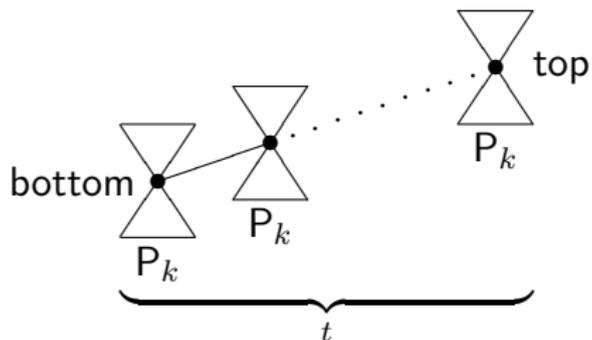
$$D(h) = 2D(h-1) = 2^h; \quad D(0) = 1$$

Beweis von d):

Anzahl der Kanten, die entfernt werden müssen:

$$\left. \begin{array}{l} C_u(h) \leq 2 + 2C_u(h-1) \\ \quad = 2 + 4 + 2^3 + \dots + 2^h \\ \quad = 2^{h+1} - 2 \\ C_d(h) \leq 1 + 2C_d(h-1) \\ \quad = 2^h - 1 \end{array} \right\} C_H \leq 2^{h+1} - 2 + 2^h - 1 \approx 3 \cdot 2^h$$

Kette von P_k 's:



Gesamtzahl der Elemente:

n

$t(2k + 1)$ in den P_k 's

$r = n - t(2k + 1)$ Rest

Wenn $r < t - 1$, dann wissen wir, dass **top** größer ist als

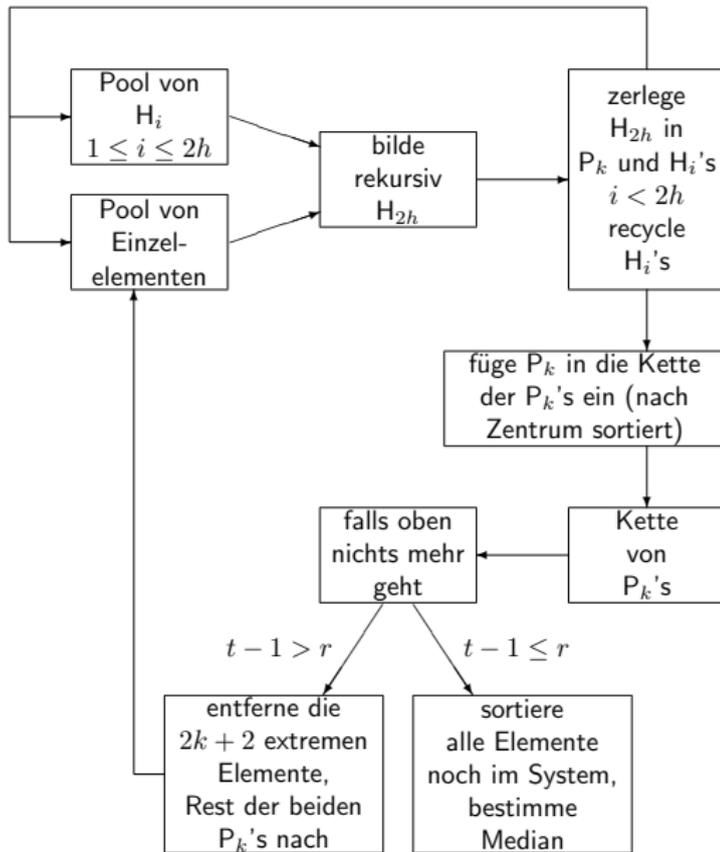
$$k + (t - 1)(k + 1) > k + (k + 1) \left(\frac{n + 1}{2k + 2} - 1 \right) = \frac{n - 1}{2}$$

$\Rightarrow \text{top} > \text{Median}$

Setze

$$k := \left\lfloor n^{\frac{1}{4}} \right\rfloor$$

$$h \text{ s.d.g. } 2^{h-1} \leq k < 2^h$$



Definiere: $r :=$ Anzahl der noch im H_{2h} -Produktionsprozess steckenden Elemente (für jedes $i < 2h$ höchstens ein H_i , daher

$$r \leq \sum_{i=0}^{2h-1} 1 = 2^{2h} - 1).$$

$R :=$ Anzahl der im letzten Schritt zu sortierenden Elemente. Es gilt: $t \leq r + 1$, und damit

$$R = t(2k + 1) + r \leq 2^{2h}(2k + 1) + 2^{2h} - 1.$$

$m :=$ Gesamtzahl der im Algorithmus produzierten P_k 's.

$$m = t + 2 \frac{n - R}{2(k + 1)} = t + \frac{n - R}{k + 1}.$$

Gesamtzahl der vom Algorithmus durchgeführten Vergleiche =

- ① Anzahl der Kanten in allen P_k 's
- ② + Anzahl der Kanten, die gelöscht werden, um die P_k 's zu formen
- ③ + Anzahl der Kanten, die zum Schluss in übriggebliebenen H_i 's, $i < 2h$, stecken
- ④ + Anzahl der Vergleiche, um jedes Zentrum der P_k 's in die (sortierte) Kette einzufügen
- ⑤ + Anzahl der Vergleiche, um die zum Schluss übriggebliebenen R Elemente zu sortieren

$$\leq \left(\frac{n-R}{k+1} + t \right) \left[\underbrace{2k}_1 + \underbrace{3k+2h}_2 + \underbrace{\log \frac{n}{2k+1}}_4 \right] + \underbrace{R \log R}_5 + \underbrace{r}_3 .$$

Mit $k = \left\lfloor n^{\frac{1}{4}} \right\rfloor$, h so, dass $2^{h-1} \leq k < 2^h$ ergibt sich damit

$$r = \mathcal{O}(k^2)$$

$$t = \mathcal{O}(k^2) \text{ zum Schluss}$$

$$R = \mathcal{O}(k^3), \text{ und damit die}$$

$$\text{Anzahl der Vergleiche} = T(n) \leq 5n + o(n).$$

Verbesserte Version (besseres Zurechtschneiden, bessere Verwertung der Reste):

$$T(n) = 3n + o(n)$$

Bester bekannter Algorithmus (von Dor/Zwick):

$$2,95n + o(n)$$

Literatur:



Arnold Schönhage, Michael Paterson, Nicholas Pippenger:
Finding the median

J. Comput. Syst. Sci. **13**, pp. 184–199 (1976)



Dorit Dor, Uri Zwick:
Selecting the median

SIAM J. Comput. **28**(5), pp. 1722–1758 (1999)