

K8. Median-of-Medians 思想 für Lösung des Auswahlproblems in erwarteter linear Zeit

Beobachtung: Wählt man Pivot-Element  $x$  in Linearzeit so, dass

$$\max \{ |A_{<x}|, |A_{>x}| \} \leq \lfloor d \cdot |A| \rfloor$$

für eine konstante  $d < 1$ , so hat Auswahl  $(A, i)$  Laufzeit  $O(n)$

Beweis: Für Laufzeitfunktion  $f: \mathbb{N} \rightarrow \mathbb{N}$  gilt dann

$$f(n) \leq f(\lfloor d \cdot n \rfloor) + \tilde{c} \cdot n$$

$d \cdot n$ : Subproblemgröße

$f(\lfloor d \cdot n \rfloor)$ : Rekursiver Aufruf, d.h. das Problem mit Größe  $\lfloor d \cdot n \rfloor$  zu lösen

$\tilde{c} \cdot n$ : Partition cost, Aufwandskosten

Warum sind die Aufwandskosten  $\tilde{c} \cdot n$ ?

Wenn wir ein Pivot  $x$  wählen, müssen wir die Menge in drei Teile

$$A_{<x}, A_{=x}, A_{>x}$$

Wir müssen das gesamte Array durchsuchen, für jedes Element eine Vergleichsoperation durchführen, um zu entscheiden, ob es links oder rechts von  $x$  liegt. Jedes Element wird höchstens einmal mit  $x$  verglichen.

Die Kosten sind  $O(n)$ .

Für  $f(n)$  und Laufzeit, gibt es zwei Beweismethoden:

Methoden-1:  $f(n) \leq f(dn) + cn$

$$\leq (f(d^2n) + c \cdot (dn)) + cn \quad \text{mit } n = dn$$

$$\leq (f(d^3n) + c \cdot (d^2n)) + c \cdot (dn) + cn$$

...

$$\leq f(d^k n) + cn \cdot (d^0 + d^1 + d^2 + \dots + d^{k-1})$$

$$\text{weil } d < 1 \Rightarrow d^0 + d^1 + \dots + d^{k-1} = \sum_{j=0}^{k-1} d^j = \frac{1-d^k}{1-d}$$

$$\Rightarrow f(n) \leq cn \cdot \frac{1}{1-d} + f(d^k n)$$

$$\text{weil } d^k \rightarrow 0, k \rightarrow \infty$$

$$\Rightarrow f(n) \leq cn \cdot \frac{1}{1-d} + t, t \in \mathbb{R}$$

$$\Rightarrow f(n) \in O(n)$$

方法 2:

use verallgemeinerter Aufteilungs - Beschleunigungs - Satz

Es sei:  $f: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  monoton wachsend,  $a \in \mathbb{R}_{>1}$ ,  $b, c \in \mathbb{R}_{\geq 0}$ ,  
so dass  $f(1) \leq \frac{c}{a}$ , und die folgende Ungleichung gilt:

$$f(a \cdot n) \leq b \cdot f(n) + c \cdot n, \text{ für alle } n \geq 1$$

Dann gilt:

$$f(n) = \begin{cases} O(n) & , a > b \\ O(n \log n) & , a = b \\ O(n^{\log_a b}) & , a < b \end{cases}$$

对于这个问题是:

$$f: \mathbb{N} \rightarrow \mathbb{N}, \text{ dann gilt } f(n) \leq f(\lfloor d \cdot n \rfloor) + \tilde{c} n$$

我们把公式放一起比较

$$f(n) \leq f(\lfloor d \cdot n \rfloor) + \tilde{c} n$$

$$f(a \cdot n) \leq b \cdot f(n) + c n$$

$$\Rightarrow b=1 \quad \lfloor d \cdot n \rfloor \cdot a = a n$$

$$\Rightarrow a = \left\lceil \frac{1}{d} \right\rceil$$

$$\text{denn } d < 1 \Rightarrow a > 1, \Rightarrow a > b$$

$$\Rightarrow f(n) \in O(n)$$

**Satz:** Wählt man in Algorithmus Auswahl  $(A, i)$  das Pivot-Element  $x \in A$  zufällig und gleichverteilt, so ist erwartete Laufzeit  $O(n)$   
证明见课件