Kapitel III Selektieren und Sortieren

1. Einleitung

Gegeben: Menge S von n Elementen aus einem total geordneten Universum U, $i \in \mathbb{N}$, $1 \leq i \leq n$.
Gesucht: i-kleinstes Element in S.

Die Fälle $i = 1$ bzw. $i = n$ entsprechen der Suche nach dem Minimum bzw. Maximum.
Der Standardalgorithmus dafür benötigt $n - 1$ Vergleiche.

Satz 76

Die Bestimmung des Minimums/Maximums von n Elementen benötigt mindestens $n-1$ Vergleiche.

Beweis:

Interpretiere Algorithmus als Turnier. Ein Spiel wird jeweils vom kleineren Element gewonnen. Wir beobachten: Jedes Element außer dem Gesamtsieger muss mindestens ein Spiel verloren haben $\Rightarrow n-1$ Vergleiche notwendig. $\square$

Bestimmung des Vize-Meisters bzw. des zweitkleinsten Elements

Satz 77

Das zweitkleinste von n Elementen kann mit

$$n + \lceil \log_2 n \rceil - 2$$

Vergleichen bestimmt werden.

Beweis:

Wir betrachten wiederum ein KO-Turnier: $(n-1)$ Vergleiche genügen zur Bestimmung des Siegers (Minimum).
Das zweitkleinste Element ist unter den „Verlierern“ gegen das Minimum zu suchen. Deren Anzahl ist $\leq \lceil \log_2 n \rceil$. Man bestimme nun unter diesen Elementen wiederum das Minimum und erhält damit das zweitkleinste Element in $\leq \lceil \log_2 n \rceil - 1$ weiteren Vergleichen. $\square$

Lewis Carroll:
Lawn Tennis Tournaments
St. Jones Gazette (Aug. 1, 1883), pp. 5–6
Reprinted in The Complete Work of Lewis Carroll. Modern Library, New York (1947)

Vaughan R. Pratt, Frances F. Yao:
On lower bounds for computing the i-th largest element
Proc. 14th Ann. IEEE SWAT, pp. 70–81 (1973)

Donald E. Knuth:
The art of computer programming. Vol. 3: Sorting and searching,
3. Auflage, Addison-Wesley Publishing Company: Reading (MA), 1997

2. Der Blum-Floyd-Pratt-Rivest-Tarjan Selektions-Algorithmus

Definition 78

Sei $n \in \mathbb{N}$. Der Median (das „mittlere“ Element) einer total geordneten Menge von n Elementen ist deren i-kleinstes Element, wobei

$$i = \left\lceil \frac{n}{2} \right\rceil .$$

Bemerkung: Für gerade n wird manchmal auch $i = \left\lfloor \frac{n}{2} \right\rfloor + 1$ benutzt.

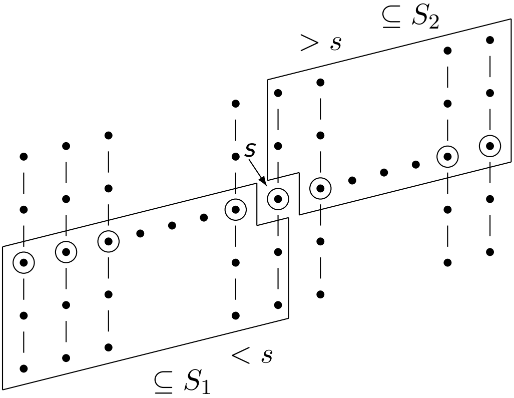
Sei m eine kleine ungerade Zahl (etwa $5 \leq m \leq 21$). Sei $S := \{a_1, \ldots, a_n\}$ eine Menge von n paarweise verschiedenen Elementen. Zur Bestimmung des i-kleinsten Element in S betrachten wir folgenden Algorithmus BFPRT.

Der BFPRT-Selektions-Algorithmus (1/3)

1. Teile S in $\left\lceil \frac{n}{m} \right\rceil$ Blöcke auf, $\left\lfloor \frac{n}{m} \right\rfloor$ davon mit je m Elementen
2. Sortiere jeden dieser Blöcke
3. Sei S' die Menge der $\left\lceil \frac{n}{m} \right\rceil$ Mediane der Blöcke. Bestimme rekursiv den Median s dieser Mediane (also das $\left\lceil \frac{|S'|}{2} \right\rceil$-kleinste Element von S').

Der BFPRT-Selektions-Algorithmus (2/3)

4. Partitioniere $S - \{s\}$ in $S_1 := \{x \in S : x < s\}$, $S_2 := \{x \in S : x > s\}$. Bemerkung: $|S_1|,\ |S_2| \geq \frac{n}{4}$, falls $n \geq m^2 + 4m$.



Der BFPRT-Selektions-Algorithmus (3/3)

5. Falls $i \leq |S_1|$, bestimme rekursiv das i-kleinste Element in S_1.
   Falls $i = |S_1| + 1$, gib s als Lösung zurück.
   Ansonsten bestimme rekursiv das $(i - |S_1| - 1)$-kleinste Element in S_2.

Sei $T(n)$ die worst-case Anzahl von Vergleichen für $|S| = n$ des Algorithmus BFPRT. Sei C_m die # von Vergleichen, um m Elemente zu sortieren (z.B. $C_5 = 7$, $C_{11} = 26$). Es gilt:

$$T(n) \leq \underbrace{T\left(\left\lceil \frac{n}{m} \right\rceil\right)}_{3.} + \underbrace{T\left(\left\lfloor \frac{3}{4} n \right\rfloor\right)}_{5.} + \underbrace{\left\lceil \frac{n}{m} \right\rceil}_{2.} C_m + \underbrace{\left\lfloor \frac{n}{2} \right\rfloor}_{4.}$$

Satz 79

Der Selektions-Algorithmus BFPRT bestimmt das i-kleinste Element von n Elementen mit $\mathcal{O}(n)$ Vergleichen (und Zeit).