

# Theoretische Informatik 1 / Algorithmentheorie

Wintersemester 2012/13

Prof. Dr. Georg Schnitger,  
Dipl. Inf. Bert Besser,  
Dipl. Inf. Matthias Poloczek

Arbeitsgruppe Theoretische Informatik, Institut für Informatik

---



## Blatt 3

Ausgabe: 01.11.2012  
Abgabe: 08.11.2012 **vor** der Vorlesung

### 3.1. Aufgabe (5+3)

*Vergleichen*

Wir betrachten das vergleichsbasierte Sortierverfahren Insertion-Sort.

- Gib den Vergleichsbaum für Insertion-Sort auf dem Eingabearray  $x_1, x_2, x_3$  an.
- In der Vorlesung haben wir Beispiele dafür gesehen, dass Bubble-Sort zwei Elemente mehrmals miteinander vergleichen kann. Gilt das auch für Insertion-Sort? Warum?

### 3.2. Aufgabe (8)

*Radix-Sort*

**Illustriere**, wie Radix-Sort das Eingabearray  $A = [69, 30, 93, 102, 23, 89, 85]$  sortiert. Dabei soll  $b = 5$  als Basis gewählt werden. **Zeige**, in  $b$ -ärer Darstellung, nach jeder Verteilphase den Inhalt aller Queues und nach jeder Sammelphase das veränderte Eingabearray.

### 3.3. Aufgabe (4+4)

*Externes Multiway-Mergesort*

Wir wollen  $n$  Datensätze sortieren, die nicht alle gemeinsam in den Hauptspeicher passen. Wir nehmen an, dass in *einem* I/O-Zugriff ein Block von genau  $B$  Datensätzen vom Externspeicher gelesen bzw. auf den Externspeicher geschrieben werden kann. Weiterhin sei  $M$  die Anzahl von Datensätzen, die sich gleichzeitig im Hauptspeicher befinden können. Der Einfachheit halber gelte  $n \equiv 0 \pmod{M}$  sowie  $M \equiv 0 \pmod{B}$ . Wir interessieren uns hier nur für die asymptotische Anzahl der I/O-Zugriffe.

- Vorbereitend wollen wir erforschen, wie  $r > 2$  sortierte Folgen schnell zu einer sortierten Folge zusammengemischt werden können. **Beschreibe** einen Algorithmus, der  $r$  sortierte Folgen mit jeweils  $k$  Schlüsseln in Zeit  $O(r \cdot k \cdot \log_2(r))$  mischt.

*Hinweis:* Heap.

- Zurück zu unserer Zielstellung, die Anzahl der I/O Zugriffe zu minimieren. Was macht ein Profi? In der ersten Phase wird sie  $n/M$  sortierte Folgen der Länge  $M$  herstellen, indem sie wiederholt den Hauptspeicher füllt und dann den gesamten Inhalt des Hauptspeichers in Form einer sortierten Folge "ausspuckt". In der zweiten Phase werden pro Schritt je-

weils  $M/B$  Folgen gleichzeitig gemischt, wobei von jeder Folge stets nur ein Block im Hauptspeicher gehalten wird - die Folgen werden "gestreamt".

Was ist die Anzahl der I/O-Zugriffe in der ersten Phase? Wieviele Schritte werden in der zweiten Phase durchgeführt? Welche Folgenlängen werden pro Schritt verarbeitet. Was ist die Gesamtanzahl der I/O-Zugriffe?