6. Aufgabenblatt zur Vorlesung

Algorithmen und Datenstrukturen — SoSe 2025

Wolfgang Mulzer

Abgabe am 06. Juni 2025 bis 10 Uhr im Whiteboard

Aufgabe 1 Elementare Wahrscheinlichkeitsrechnung *15 Punkte*

Auf einem Tisch stehen N Kisten. In diese Kisten werden nacheinander unabhängig voneinander n Bälle geworfen, wobei jede Kiste mit gleicher Wahrscheinlichkeit getroffen wird.

(a) Berechnen Sie die Wahrscheinlichkeit, dass Kiste i leer ist. Sei Y_i die Zufallsvariable, die den Wert 1 annimmt, falls Kiste i leer ist, und 0 sonst. Geben Sie auch den Erwartungswert $\mathbf{E}[Y_i]$ an.

Hinweis: Wenn Kiste i leer bleibt, dann landen alle n Bälle in den $N-1$ anderen Kisten.

(b) Sei X die Zufallsvariable, welche die Anzahl von leeren Kisten angibt. Berechnen Sie den Erwartungswert von X mit Hilfe der Erwartungswerte $\mathbf{E}[Y_i]$.

(c) Bestimmen Sie, wie viele Bälle man benötigt, damit gilt: mit Wahrscheinlichkeit mindestens $1/2$ enthält mindestens eine Kiste mindestens zwei Bälle.

Hinweis: Stellen Sie sich vor, die n Bälle werden nacheinander in die Kisten geworfen. Was ist die Wahrscheinlichkeit, dass der nächste Ball in einer leeren Kiste landet, wenn alle vorherigen Bälle in einer leeren Kiste gelandet sind? Verwenden Sie die ungemein nützliche Abschätzung $1+x \leq e^x$, welche für alle $x \in \mathbb{R}$ gilt.

(d) Professor Pinocchio hat eine Idee, um Hashtabellen zu vereinfachen. Wenn wir die Zahl N der Plätze in der Hashtabelle im Verhältnis zur Anzahl der zu speichernden Einträge n groß genug wählen, sollte die Wahrscheinlichkeit, dass Kollisionen auftreten, verschwindend gering werden (unter der Annahme, dass sich die Hashfunktion wie eine zufällige Funktion verhält). Dann könnte man auf die Kollisionsbehandlung verzichten. In Anbetracht von (c), was halten Sie von dem Vorschlag? (Wenn Sie Teil (c) nicht gelöst haben, dann bearbeiten Sie diesen Teil unter der Annahme, dass für $N^{1/3}$ Bälle mindestens eine Kollision auftritt.

Aufgabe 2 Hashing: Worst-case-Analyse *10 Punkte*

Sei A eine Hashtabelle der Größe N, und $n \in \mathbb{N}$ beliebig. Zeigen Sie: Für jede Schlüsselmenge K mit $|K| \geq (n-1)N+1$ und jede Hashfunktion $h : K \to \{0, \ldots, N-1\}$ existiert eine Menge $S \subseteq K$ mit $|S| = n$, so dass alle Elemente von S auf denselben Eintrag in A abgebildet werden. Was bedeutet das für die worst-case Laufzeit von Hashing mit Verkettung? Wie verträgt sich das mit der Analyse aus der Vorlesung?

Aufgabe 3 Hashing im Selbstversuch *5 Punkte*

Fügen Sie nacheinander die Schlüssel 5, 28, 19, 15, 20, 33, 12, 17, 10 in eine Hashtabelle der Größe 9 ein. Die Hashfunktion sei $h(k) = k \bmod 9$. Sie mit Die Konflikte werden mit Verkettung gelöst.

Illustrieren Sie jeweils die einzelnen Schritte.

Wiederholen Sie dann mit der Hashfunktion $h(k) = k \bmod 6$ und einer Hashtabelle der Größe 6.