Rev. 1, 23.04.2018
TU Ilmenau, Fakultät für Informatik und Automatisierung
FG Komplexitätstheorie und Effiziente Algorithmen
Univ.-Prof. Dr. M. Dietzfelbinger, M. Sc. P. Schlag
http://www.tu-ilmenau.de/iti/lehre/lehre-ss-2018/ra/

TECHNISCHE UNIVERSITÄT
ILMENAU

Randomisierte Algorithmen SS 2018 – Übung 3

Besprechung: Montag, 14. Mai 2018

Hinweis: Für das erfolgreiche Vorrechnen einer mit „*“ gekennzeichneten Aufgabe wird ein Bonuspunkt vergeben, es gibt maximal zwei Bonuspunkte pro Studierendem im Semester. Bitte geben Sie Ihren Lösungsvorschlag bis zum Vortag der Übung (Sonntag), 13:00 Uhr, per E-Mail an philipp.schlag@tu-ilmenau.de oder bis Mittwoch, 17:00 Uhr, direkt in meinem Büro (Z 1048) ab.

Aufgabe 1 (Maximale Last beim vollständig zufälligen Hashing) *

Wir betrachten das Zufallsexperiment, n Bälle auf m Körbe zu werfen (äquivalent: n Schlüssel mit einer rein zufälligen Hashfunktion auf m Fächer zu verteilen), wobei jeder Korb bei einem einzelnen Wurf mit gleicher Wahrscheinlichkeit $1/m$ getroffen wird. Es sei $\alpha = n/m$ der Auslastungsfaktor mit $\alpha \leq \alpha_0$ für eine Konstante $\alpha_0 > 0$.

(a) Geben Sie einen geeigneten Wahrscheinlichkeitsraum (Ω, p) an.

(b) Sei $k \in \mathbb{N}$ und $j \in \{1, \ldots, m\}$ fest. Es sei $p_{k,j}$ die Wahrscheinlichkeit dafür, dass das Ereignis $\mathscr{E}_{k,j} :=$ „Korb Nummer j enthält mindestens k Bälle“ eintritt. Beweisen Sie:

$$p_{k,j} = \mathbf{Pr}(\mathscr{E}_{k,j}) \leq \frac{\alpha^k}{k!}.$$

Hinweis: Wenn Korb j mindestens k Bälle enthält, dann muss es eine Menge $A \subseteq \{1, \ldots, n\}$ mit $|A| = k$ geben, so dass alle Bälle mit Nummer $i \in A$ in Korb j liegen.

(c) Sei $k \in \mathbb{N}$ fest. Geben Sie eine obere Schranke $u(n,k)$ für die erwartete Anzahl der Körbe mit mindestens k Bällen an.

Hinweis: Betrachten Sie die Zufallsvariable $X_k :=$ „Anzahl der Körbe mit mindestens k Bällen“.

(d) Zeigen Sie:

$$\mathbf{Pr}(\text{„es gibt einen Korb mit mindestens } k \text{ Bällen“}) \leq n \cdot \frac{\alpha^{k-1}}{k!}.$$

Hinweis: Nutzen Sie die Markov-Ungleichung.

(e) Zeigen Sie: Gilt $|\ln \alpha_0| + \ln \alpha_0 + 1 < \ln k - \frac{\ln n}{k}$ für $k \geq 1$, so folgt $n \cdot \frac{\alpha^{k-1}}{k!} < 1$.

Hinweis: Wegen $\frac{k^k}{k!} \leq \sum_{i=0}^{\infty} \frac{k^i}{i!} = e^k$ gilt $k! \geq \left(\frac{k}{e}\right)^k$.

(f) Zeigen Sie, dass für die Wahl $k = k_2 := \frac{2 \ln n}{\ln \ln n}$ und genügend große n gilt:

$$\mathbf{Pr}(\text{„es gibt einen Korb mit mindestens } k \text{ Bällen“}) < 1.$$

Weiter auf nächster Seite!

(g) Sei L die maximale Beladung eines Korbes. Zeigen Sie, dass für genügend große n gilt:

$$\mathbf{E}(L) \leq k_2 + O(1).$$

Hinweis: Nützlich sind Fakt 2.2.8 sowie folgende Beobachtung:
Sei $a_i = n \cdot \frac{\alpha^{i-1}}{i!}$. Für $i \geq k_2$ gilt $\frac{a_{i+1}}{a_i} = \frac{\alpha}{i+1} \leq \frac{\alpha}{k_2+1}$. Für genügend große n gilt $\frac{\alpha}{k_2+1} \leq \beta < 1$ für eine Konstante β.

Aufgabe 2 (Tail bounds) *

Es seien $X_1, \ldots, X_n$ unabhängige Zufallsvariable mit $\mathbf{Pr}(X_i = 1) = \mathbf{Pr}(X_i = -1) = \frac{1}{2}$. Sei $X = X_1 + \cdots + X_n$. Offensichtlich ist $\mathbf{E}(X) = 0$. Wir wollen untersuchen, mit welcher Wahrscheinlichkeit X weit von diesem Mittelwert entfernt ist, d. h. wir wollen $\mathbf{Pr}(|X| \geq a)$ für $0 < a \leq n$ abschätzen. Dazu nutzen wir aus der Vorlesung bekannte Ungleichungen und Schranken.

(a) **Chebychev-Ungleichung.** Berechnen Sie $\mathbf{Var}(X)$ und benutzen Sie die Chebychev-Ungleichung (Fakt 2.3.3), um eine Abschätzung der Form $\mathbf{Pr}(|X| \geq a) \leq \ldots$ zu erhalten. Für welche Werte von a ergibt sich eine nicht-triviale obere Schranke (d. h. ein Wert kleiner als 1)?

(b) **Chebychev-Cantelli-Ungleichung.** Gehen Sie wie in (a) vor, verwenden Sie allerdings die Chebychev-Cantelli-Ungleichung (Proposition 2.7.2). Für welche Werte von a ergibt sich eine nicht-triviale obere Schranke? Für welche Werte von a ist diese besser als die Schranke aus (a)?

(c) **Hoeffding-Ungleichung.** Vergewissern Sie sich zunächst, dass Sie die Hoeffding-Ungleichung auf $\frac{1}{2}(X_1+1), \ldots, \frac{1}{2}(X_n+1)$ anwenden dürfen (siehe Satz 2.6.1). Wählen Sie anschließend die Werte m und ε passend, um mit Hilfe von Korollar 2.6.3 eine obere Schranke für $\mathbf{Pr}(|X| \geq a)$ zu erhalten.

(d) **Chernoff-Schranke.** Nun wollen wir die in den Beispielen zur **verallgemeinerten Markov-Ungleichung** (Proposition 2.3.4) beschriebene Chernoff-Schranke nutzen. Sei dazu $t > 0$ beliebig.

- **(i)** Zeigen Sie: $\mathbf{E}(e^{tX_i}) < e^{t^2/2}$.
  Hinweis: Man kann den Erwartungswert explizit hinschreiben, die Taylorreihen der beiden Summanden und von $e^{t^2/2}$ ermitteln und abschätzen.
- **(ii)** Folgern Sie aus (i): $\mathbf{E}(e^{tX}) < e^{t^2 n/2}$.
- **(iii)** Folgern Sie aus (ii): $\mathbf{Pr}(X \geq a) \leq e^{t^2 n/2 - ta}$.
- **(iv)** Finden Sie zu gegebenem a den Wert t, der $e^{t^2 n/2 - ta}$ minimiert. Folgern Sie eine zugehörige obere Schranke für $\mathbf{Pr}(X \geq a)$. [Resultat: $e^{-a^2/(2n)}$.]
- **(v)** Was kann man über $\mathbf{Pr}(X \leq -a)$ und über $\mathbf{Pr}(|X| \geq a)$ sagen? Formulieren Sie eine Aussage und beweisen Sie sie.
- **(vi)** Für welche Werte von a ergibt sich eine nicht-triviale obere Schranke für $\mathbf{Pr}(|X| \geq a)$?

(e) Vergleichen Sie Ihre Ergebnisse aus (a) – (d). (Wählen Sie beispielsweise ein festes n und stellen Sie die oberen Schranken für $0 < a \leq n$ gemeinsam graphisch dar.)