Ausarbeitung zur Vorlesung

Algorithmen, Datenstrukturen und Datenabstraktion WS 2019/20

Günter Rote

Erwartete Höhe einer Skipliste

Im Folgenden soll die erwartete Höhe einer Skipliste mit n Elementen abgeschätzt werden. Untersucht wird das Zufallsexperiment des Aufbaus einer Skipliste mit n Elementen. Sei H eine Zufallsvariable mit $H =$ Höhe der Skipliste. Gesucht ist eine Abschätzung für $E[H]$.

Der Erwartungswert ist wie folgt definiert:

$$E[H] = \sum_{i=1}^{\infty} \Pr[H = i] \cdot i$$

Diese Summe kann man auch so schreiben:

$$\begin{array}{llllllll}
 & \Pr[H=1] & & & & & & \\
+ & \Pr[H=2] & + & \Pr[H=2] & & & & \\
+ & \Pr[H=3] & + & \Pr[H=3] & + & \Pr[H=3] & & \\
+ & \Pr[H=4] & + & \Pr[H=4] & + & \Pr[H=4] & + & \Pr[H=4] \\
+ & \cdots & & & & & &
\end{array}$$

Wenn die Summe nun spaltenweise gelesen wird, ergibt sich:

$$E[H] = \sum_{i=1}^{\infty} \Pr[H = i] \cdot i = \sum_{i=1}^{\infty} \sum_{j=i}^{\infty} \Pr[H = i] = \sum_{i=1}^{\infty} \Pr[H \geq i] = \sum_{i=1}^{\infty} \Pr[H > i - 1]$$

In Aufgabe 34 wird gezeigt, dass $\Pr[H > k] \leq \frac{n}{2^k}$ ist. Außerdem ist jede Wahrscheinlichkeit höchstens 1. Die ersten $\lceil \log n \rceil$ Summanden werden daher mit 1 abgeschätzt, alle weiteren mit $\frac{n}{2^k}$.

$$\sum_{i=1}^{\infty} \Pr[H > i - 1] \leq \sum_{i=1}^{\lceil \log n \rceil} 1 + \sum_{i=\lceil \log n \rceil + 1}^{\infty} \frac{n}{2^{i-1}} = \lceil \log n \rceil + \sum_{i=\lceil \log n \rceil + 1}^{\infty} \frac{n}{2^{i-1}}$$

Es muss also noch die verbliebene Summe abgeschätzt werden. Der erste Summand ist $n/2^{\lceil \log n \rceil} \leq 1$. Jeder weitere ist immer halb so groß wie der Vorangegangene. Die Gesamtsumme ist also kleiner oder gleich der geometrischen Reihe $1 + \frac{1}{2} + \frac{1}{4} + \cdots = 2$:

$$\lceil \log n \rceil + \sum_{i=\lceil \log n \rceil + 1}^{\infty} \frac{n}{2^{i-1}} \leq \lceil \log n \rceil + \sum_{i=0}^{\infty} \frac{1}{2^i} = \lceil \log n \rceil + 2$$

Die erwartet Höhe H wächst also logarithmisch mit der Anzahl der Elemente.