

Algorithmen und Datenstrukturen (Th. Ottmann und P. Widmayer)

Folien: Analyse Ideales Hashing

Autor: Stefan Edelkamp

Institut für Informatik
Georges-Köhler-Allee
Albert-Ludwigs-Universität Freiburg

1 Ideales Hashing

Beim **idealen Hashing** wird angenommen, dass n Daten in einem Speicher mit m Plätzen eingefügt werden, keine Löschungen vorgenommen worden sind und alle Konfigurationen von n besetzten und $m - n$ unbesetzten Speicherplätzen die gleiche Wahrscheinlichkeit $1 / \binom{m}{n}$ besitzen. Sei P_r die Wahrscheinlichkeit, dass beim der erfolglosen Suche genau r Plätze getestet werden müssen, dann gilt:

$$P_r = \begin{cases} \binom{m-r}{n-(r-1)} / \binom{m}{n} & \text{für } 1 \leq r \leq m \\ 0 & r > m \end{cases}$$

Begründung: Die ersten $r - 1$ Plätze sind besetzt, der r -te Platz ist frei. Auf den restlichen $m - r$ Plätzen können die weiteren $n - (r - 1)$ besetzten Plätze beliebig verteilt sein.

Erwartete Anzahl von Tests

$$\begin{aligned}
 C'_n &= \sum_{r=1}^m r P_r = \sum_{r=1}^m r \binom{m-r}{n-(r-1)} / \binom{m}{n} \\
 &= \sum_{r=1}^m r \binom{m-r}{m-n-1} / \binom{m}{n} \\
 &= \sum_{r=1}^m ((m+1) - (m+1-r)) \binom{m-r}{m-n-1} / \binom{m}{n} \\
 &= (m+1) - \sum_{r=1}^m (m+1-r) \binom{m-r}{m-n-1} / \binom{m}{n} \\
 &= (m+1) - \sum_{r=1}^m (m-n) \binom{m-r+1}{m-n} / \binom{m}{n} \\
 &= (m+1) - (m-n) / \binom{m}{n} \sum_{r=1}^m \binom{m-r+1}{m-n} \\
 &= (m+1) - (m-n) / \binom{m}{n} \sum_{r=1}^m \binom{r}{m-n} \\
 &= (m+1) - (m-n) / \binom{m}{n} \binom{m+1}{m-n+1} \\
 &= (m+1) - (m-n) / \binom{m}{n} \binom{m+1}{n} \\
 &= (m+1) - (m-n)(m+1) / (m-n+1) \\
 &= (m+1)(1 - (m-n)/(m-n+1)) \\
 &= (m+1) / (m-n+1) \approx m / (m-n) \\
 &= 1 / (1 - \alpha) = 1 + \alpha + \alpha^2 + \alpha^3 + \dots
 \end{aligned}$$

Durchschnittliche Anzahl Inspizierter Plätze

$$\begin{aligned}C_n &= \frac{1}{n} \sum_{k=0}^{n-1} (m+1)/(m-k+1) \\&= \frac{1}{n} \sum_{k=1}^n (m+1)/(m-k+2) \\&= \frac{m+1}{n} \left(\frac{1}{m-n+2} + \frac{1}{m-n+3} + \dots + \frac{1}{m-1} \right) \\&= \frac{m+1}{n} (H(m+1) + H(m-n+1)) \\&\approx \frac{m+1}{n} \ln \frac{m+1}{m-n+1} \approx \frac{1}{\alpha} \ln \frac{1}{1-\alpha}\end{aligned}$$

mit $H(m) = 1 + 1/2 + 1/3 + \dots + 1/m \approx \ln m$.