

Übungen *Algorithm Engineering* - Blatt 5

Abgabe 14.6.2010, Besprechung 17.6.2010

1. (Allgemeine Perfekte Hashfunktionen) Zeigen Sie in zwei Schritten, dass der Speicherplatz für das perfekte Hashen von n Elementen aus einem Universum mit u Elementen in eine Tabelle der Größe m mindestens

$$\log_2 e \cdot \left(\sum_{i=1}^n \ln(1 - i/u) - \ln(1 - i/m) \right)$$

Bits beträgt.

- (a) Zeigen Sie zuerst, dass für die Menge der perfekten Hashfunktionen auf n Elementen gilt: $|H| \geq \frac{\binom{u}{n}}{(u/m)^n \binom{m}{n}}$.

- (b) Leiten Sie daraus ab, dass

$$\log_2 |H| \geq \left(\sum_{i=1}^n \log_2(1 - i/u) - \log_2(1 - i/m) \right)$$

gilt.

2. (Spezielle Perfekte Hashfunktionen)

Der Zauberwürfel besitzt $n = 8! \cdot 3^8 \cdot 12! \cdot 2^{12} / (3 \cdot 2 \cdot 2) = 43\,252\,003\,274\,489\,856\,000$ viele Zustände (in Bits entspricht das 4.73 Exabytes).

- (a) Bestimmen Sie eine in $O(n)$ zu berechnende, perfekte Hashfunktion (rank), d.h., eine injektive Abbildung der Würfelzustände in $\{0, \dots, 2n - 1\}$.
- (b) Bestimmen Sie eine in linearer Zeit zu berechnende Umkehrfunktion (unrank)
- (c) * Bestimmen Sie eine in $O(n)$ zu berechnende, minimale perfekte Hashfunktion bzw. bijektive Abbildung der Würfelzustände in $\{0, \dots, n - 1\}$.

3. (Approximatives Matching)

Bestimmen Sie den

- (a) Abhängigkeitsgraphen für die approximative Zeichenkettensuche
- (b) Spurgraphen für die Editierdistanz

bzgl. $A = (1, 2, 2, 4, 1, 4, 3, 2, 3, 2, 3)$ und $B = (1, 4, 2, 2, 3)$ über dem Alphabet $\Sigma = \{1, 2, 3, 4\}$ mit den Kosten von 1 für das Einfügen, Löschen und Ersetzen eines Zeichens.