Technische Universität München
Fakultät für Informatik
Lehrstuhl für Effiziente Algorithmen
Prof. Dr. Ernst W. Mayr/Dr. Sven Kosub
Moritz Maaß

Sommersemester 2003/04
Übungsblatt 4
23. Juni 2004

Internet-Algorithmik

Abgabetermin: 16.06.2004 **vor** *der Vorlesung*

Es seien das Universum $\mathcal{U} = [1, N]$ und die Hash-Tabelle $T = [1, n]$ gegeben.
Eine Klasse $\mathcal{H}$ von Hash-Funktionen heißt *paarweise unabhängig* genau dann, wenn

$$\frac{||\{h \in \mathcal{H} | h(x) = k_1, h(y) = k_2\}||}{||\mathcal{H}||} = \frac{1}{n^2}$$

für alle $x, y \in \mathcal{U}$ mit $x \neq y$ und alle $k_1, k_2 \in T$ gilt.

Aufgabe 1

Sei p eine Primzahl und $N = n = p$. Für $a, b \in \mathbb{Z}_p$ sei $h_{a,b}$ definiert als $h_{a,b}(i) = a \cdot i + b \bmod p$. Zeigen Sie, dass die Menge $\mathcal{H} = \{h_{a,b} | a, b \in \mathbb{Z}_p\}$ eine Klasse von paarweise unabhängigen Hash-Funktionen ist.

Aufgabe 2

Sei $N = 2^k$ und $n = 2^l$. Sei M eine $k \times l$ Matrix mit Einträgen aus $\{0, 1\}$ und b ein l-dimensionaler Vektor. Für $x = \sum_{i=0}^{n-1} x_i 2^i \in \mathcal{U}$ definieren wir

$$h_M(x) = \sum_{j=0}^{l-1} 2^j \cdot \left[\left(\left(\sum_{i=0}^{k-1} M_{j,i} x_i \right) + b_j \right) \bmod 2 \right].$$

Zeigen Sie, dass die Menge

$$\mathcal{H} = \{h_{M,b} | M \text{ ist eine } k \times l \text{ Matrix}, b \text{ ist ein Vektor mit Dimension } l\}$$

eine Klasse von paarweise unabhängigen Hash-Funktionen ist.

Aufgabe 3

Sei $\mathcal{H}$ eine Klasse von paarweise unabhängigen Hash-Funktionen. Sei $S \subseteq \mathcal{U}$ eine Menge von Schlüsseln. Sei $\alpha = \frac{|S|}{n}$. Zeigen Sie, dass

$$\left(\alpha - \frac{\alpha^2}{2} \right) \leq \mathbb{P}[\text{es gibt ein } x \in S \text{ mit } h(x) = t] \leq \alpha$$

gilt, wobei die Wahrscheinlichkeit über die zufällige Auswahl von $h \in \mathcal{H}$ und $t \in T$ definiert ist.

Aufgabe 4

Schreiben Sie ein Java-Programm, welches die exponenziellen Histogramme aus der Vorlesung implementiert. Erstellen Sie dazu eine Klasse `ExponentialHistogram` mit einem Konstruktor, der als Parameter die Werte `int k, int N` erhält, wobei `k` dem Wert $\lceil 1/\epsilon \rceil$ entspricht und `N` die Fenstergröße ist. Die Klasse sollte desweiteren die öffentlichen Methoden `void nextItem(bool v)` sowie `int currentCount()` haben.
Schreiben Sie eine `main`-Methode, die die Parameter `k` und `N` als Argumente liest (`Integer.parseInt()`) und einen Datenstrom von der Standard-Eingabe ließt. Der Datenstrom besteht aus den Buchstaben `0` und `1`, die dem Histogramm übergeben werden sollen. Nach jedem Element soll der aktuelle Wert der Anzahl der Einsen ausgegeben werden.
Bemühen Sie sich um Platz- und Zeiteffizienz und schicken Sie Ihren Source-Code (und nur diesen) an `maass@in.tum.de`. Dabei sollten es sich nur um eine Datei `ExponentialHistogram.java` handeln, weitere ggf. benötigte Klassen sind als Unterklassen zu definieren.