PD Dr. rer. nat. habil. Sven Groppe

Übungen zur Vorlesung

Semantic Web

WS 2013/2014

Übung 11 – Parallele Datenbanken

Aufgabe 1:

Sei H eine endliche Menge von Hash-Funktionen mit Signatur $U \rightarrow \{ 0, 1, \ldots, m\text{-}1 \}$. H wird universell genannt, falls für jedes Paar von verschiedenen Schlüsseln $x, y \in U$ die Anzahl der Hash-Funktionen $h \in H$ mit $h(x) = h(y)$ genau $|H|/m$ ist. Mit anderen Worten, die Chance einer Kollision zwischen x und y mit $x \neq y$ ist genau $1/m$ bei einer zufällig gewählten Hash-Funktion $h \in H$, also genau so groß wie die Chance einer Kollision bei Hash-Funktionen, die alle Elemente aus U zufällig auf $\{ 0, 1, \ldots, m\text{-}1 \}$ abbilden. Bemerken Sie, dass für die letztgenannte Art von Hash-Funktionen eine Tabelle der Größe $|U|$ notwendig wäre, um die Abbildungen von U auf $\{ 0, 1, \ldots, m\text{-}1 \}$ zu verwalten.

Die folgende Klasse H_{prim} von Hash-Funktionen ist universell:

Sei m eine Primzahl. Wir zerlegen einen Schlüssel x in $r + 1$ Unterkomponenten, so dass $x = < x_0, x_1, \ldots, x_r >$. Die einzige Bedingung ist, dass der maximale Wert einer Unterkomponente kleiner als m sein muss. Sei nun weiterhin $a = < a_0, a_1, \ldots, a_r >$ eine Sequenz von $r + 1$ Elementen, die zufällig aus $\{ 0, 1, \ldots, m\text{-}1 \}$ gewählt sind. Dann sei eine entsprechende Hash-Funktion $h_a \in H_{prim}$ durch $h_a(x) = (\Sigma_{i=0}^{r} a_i * x_i) \bmod m$ definiert. Bemerken Sie, dass wir durch Zahlentheorie beweisen können, dass H_{prim} universell ist.

a) Berechnen Sie die Hash-Werte von 3, 20 und 82 bezüglich der Hash-Funktion $h_{<5, 8, 9>}(x) = (\Sigma_{i=0}^{r} a_i * x_i) \bmod 13$. Eine Unterkomponente x_i habe dabei den Wertebereich $\{0, \ldots, 7\}$ und werde durch $x_i = (x \text{ div } 8^i) \bmod 8$ berechnet.

b) Zeigen Sie: Falls wir $a_i > 0$ fordern würden für jedes $a_i \in \{ a_0, a_1, \ldots, a_r \}$, dann ist H_{prim} nicht mehr universell.