Winter Term 2001/02 | Prof. Dr. Susanne Albers
December 3, 2001 | Institute of Computer Science

Algorithms Theory

Problem Set 4

To be submitted on December 17, 2001 (in the lecture)

Problem 1: (10 points)
Consider the standard hashing scheme where collisions as simply resolved by chaining. Let $h : [0..N-1] \to [0..m-1]$ be a hash function and $S \subseteq [0..N-1]$ be a set with $|S| = n$. Suppose that the table position of each element $x \in S$ is a random position, i.e. $\text{prob}[h(x) = i] = 1/m$ for $i = 0, \dots, m-1$, independent of the position of the other elements $y \neq x$. What is the total expected number of collisions, i.e. what is the expected value of $|\{(x, y) : x, y \in S, x \neq y, h(x) = h(y)\}|$?

Problem 2: (10 points)
Let $H \subseteq \{h; h : [0..N-1] \to [0..m-1]\}$ be c-universal, and $U = [0..N-1]$. Prove that $c \geq 1 - m/N$.
Hint: Prove that $\sum_{x,y \in U} \delta_h(x, y) \geq N(N/m - 1)$ for every function $h : [0..N-1] \to [0..m-1]$. This implies that there exist x_1 and y_1 with $\sum_{h \in H} \delta_h(x_1, y_1) \geq |H|(N/m - 1)/N$.

Problem 3: (10 points)
For $k = (k_2, k_1, k_0) \in [0..N-1]^3$ let

$$h_k : [0..N-1] \longrightarrow [0..m-1]$$

be defined as $h_k(x) = ((k_2 x^2 + k_1 x + k_0) \bmod N) \bmod m$. Moreover, for $S \subseteq [0..N-1]$ with $|S| = n$, define $b_{ik} = |\{x \in S; h_k(x) = i\}|$. Prove the following relation:

$$\sum_{k \in [0..N-1]^3} \sum_{i=0}^{m-1} b_{ik}(b_{ik} - 1)(b_{ik} - 2) \leq \left(\frac{2n}{m}\right)^2 nN^3.$$

Problem 4: (10 points)
Develop a deterministic algorithm that determines a k, as specified in part 1 of Corollary 2 presented for perfect hashing, in time $O(m + nN)$.