CMPSC 623 Problem Set 4.
by Prof. Honggang Zhang

Out: Oct. 18, 2007
Due: Oct. 25, 2007, before class.

Problem 1. Page 228, Exercise 11.2-1.

Problem 2. Page 236, Exercise 11.3-3. You are not required to give an example of an application (as required in the original problem).

Problem 3. Page 244, Exercise 11.4-1. Only do double hashing (Page 240).

Solution:

((10 mod 11) + 0*(1+ (10 mod 10))) mode 11 = 10

((22 mod 11) + 0*(1+ (22 mod 10))) mode 11 = 0

((31 mod 11) + 0*(1+ (31 mod 10))) mode 11 = 9

((4 mod 11) + 0*(1+ (4 mod 10))) mode 11 = 4

((15 mod 11) + 0*(1+ (15 mod 10))) mode 11 = 4, collision

((15 mod 11) + 1*(1+ (15 mod 10))) mode 11 = 10, collision

((15 mod 11) + 2*(1+ (15 mod 10))) mode 11 = 5

((28 mod 11) + 0*(1+ (28 mod 10))) mode 11 = 6

((17 mod 11) + 0*(1+ (17 mod 10))) mode 11 = 6, collision

((17 mod 11) + 1*(1+ (17 mod 10))) mode 11 = 3

((88 mod 11) + 0*(1+ (88 mod 10))) mode 11 = 0, collision

((88 mod 11) + 1*(1+ (88 mod 10))) mode 11 = 9, collision

((88 mod 11) + 2*(1+ (88 mod 10))) mode 11 = 7

((59 mod 11) + 0*(1+ (59 mod 10))) mode 11 = 4, collision

((59 mod 11) + 1*(1+ (59 mod 10))) mode 11 = 3, collision

((59 mod 11) + 2*(1+ (59 mod 10))) mode 11 = 2

Problem 4. Page 250, Problem 11-2 (a) (b).

Solution:

(a). Note that here the table size is n. A given key is hashed to a chosen slot with probability $1/n$. For a given slot, if we choose k keys, then the probability that all of them are hashed into that slot and all other other keys are hashed to other slots is: $(1/n)^k(1-1/n)^{(n-k)}$. As there are C_k^n possibilities, thus, we get:

$$C_k^n(1/n)^k(1-1/n)^{(n-k)}.$$

(b). Let x_i be the number of keys in slot i. Then, we have:

$$\begin{aligned}
P_k &= P\{M = k\} &(1)\\
&= P\{(\max_i \quad x_i) = k\} &(2)\\
&= P\{(\exists i \quad \text{such that} \quad x_i = k) \quad \text{and} \quad (\forall i, x_i \leq k)\} &(3)\\
&\leq P\{(\exists i \quad \text{such that} \quad x_i = k)\} &(4)\\
&= P\{(x_1 = k) \quad \text{or} \quad (x_2 = k) \quad \text{or} \quad (x_3 = k) \quad \text{or} \quad ... \quad \text{or} \quad (x_n = k)\} &(5)\\
&\leq \sum_i P\{x_i = k\} &(6)\\
&= nQ_k &(7)
\end{aligned}$$

Problem 5. Suppose that a hash table has m slots, and we resolve collisions by chaining. There are n keys. Assume that our hashing is simple uniform hashing. If you randomly pick up a slot from the table, what is the probability that you find that slot is not empty? What is the probability that you find at least two keys in that slot? If you pick up a key and identify the slot where this key is in, then what is the probability that some other keys will be in this slot?

Solution:

Probability that a key is in a randomly picked slot is $q = \frac{1}{m}$, and the probability that a key is in other remaining slots is $1 - q$.

Probability that a randomly picked slot is not empty is: $1 - (1 - q)^n = 1 - (\frac{m-1}{m})^n$.

Probability that a randomly picked slot has at least two keys:

$$\begin{gathered}
1 - (1 - q)^n - n * q * (1 - q)^{(n-1)}\\
= 1 - (\tfrac{m-1}{m})^n - n * (1/m) * (\tfrac{m-1}{m})^{(n-1)}
\end{gathered}$$

If you identify the slot where a given key is in, then the probability that some other key(s) will be in this slot is $1 - ((m - 1)/m)^{n-1}$.

Problem 6. Show that if we restrict each component a_i of a in the universal hash function $h_a(k)$ discussed in class to be nozero, then the set $H = \{h_a\}$ is not universal. (Hint: consider the keys $x = 0$ and $y = 1$.)