CSE331 Homework 5

Due by the class on 10/29/2013 (Tuesday)

All the textbook problems are in the 3rd edition. You can borrow mine or TA's if you don't have it. Doing wrong problems will not earn you any credit.

1. Do textbook problem 5.1 on page 208. h(x)=x mod 10. Read 5.4.2 about quadratic probing. (10 pts)

2. Professor Marley hypothesizes that substantial performance gains can be obtained if we modify the chaining scheme so that each list is kept in sorted order. How does the professor's modification affect the running time for successful searches, unsuccessful searches, insertions, and deletions? (5 pts)

3. Do textbook problem 5.8. Specifically, compare chaining strategy, linear probing, and double hashing. (5 pts)

4. We talked about pattern search using hashing strategy in class. Suppose we want to find the first occurrence of a string $P_1P_2\ldots P_k$ in a long input string $A_1A_2\ldots A_N$. We can solve this problem by hashing the pattern string, obtaining a hash value H_p, and comparing this value with the hash value formed from $A_1A_2\ldots A_k$, $A_2A_3\ldots A_{k+1}$, $A_3A_4\ldots A_{k+2}$, and so on until $A_{N-k+1}A_{N-k+2}\ldots A_N$. If we have a match of hash values, we compare the strings character by character to verify the match. We return the position (in A) if the strings actually do match, and we continue in the unlikely event that the match is false.
   a) Design a hash function so that the following property is satisfied: if the hash value of $A_iA_{i+1}\ldots A_{i+k-1}$ is known, then the hash value of $A_{i+1}A_{i+2}\ldots A_{i+k}$ can be computed in constant time. (5 pts)
   b) Write an algorithm which uses your designed hash function in part a) and has the running time O(k+N) plus the time spent refuting false matches. (5 pts)

5. Consider a version of the division method in which h(k)=k mod m, where $m=2^p-1$ and k is a character string interpreted in radix 2^p. **Compute** h(ACG), h(CGA), and h(AGC). What did you observe from these three examples? (5 pts)