$\boxed{\text{Burt Rosenberg}}$

Problem Set 9

OUT: 9 NOVEMBER, 1993
DUE: 16 NOVEMBER, 1993

Goals

Understanding the mathematics of Hashing.

Reading Assignment

Read Chapters 16 and 17 from *Algorithms.*

Food for Thought

In class we reiterated the advice given in the text:

> *Make the size of your hash tables a prime number.*

This done, any integer between 1 and the hash table's size minus 1 can be used as a skip value. If the table size is not a prime, it is possible that most of the table will be inaccessible to double hashing's collision resolution scheme.

It is often convenient to think in terms of perfect powers of two, integers of the form 2^i, for an integer i. Consider the case of a table of size 4. If the skip is 1, then the entire table can be visited moving in steps of 1. This is also true of a skip of 3, starting from 0, successive increments of 3, taken mod 4, give the sequence,

$$0, 3, 2, 1.$$

1. Show by direct calculation that in a table of size 8, any *odd skip* will succeed in visiting every entry in the table. That is, for skip k, then $0, k, 2k, \ldots$ mod 8 gives (in some order) all the integers 0 through 7.

2. Show that this is not true for some even skip value.

3. Prove this is so for general numbers of the form 2^i.

4. Explain how the previous fact can be used to implement good double hashing schemes for hash tables of size a pure power of two.

Programming Assignment

Write a program that automatically performs the above experiment but for general table sizes and skip values. That is, the program accepts two values n and k, and beginning at $i \leftarrow 0$ repeats the update

$$i \leftarrow (i + k) \pmod{n}$$

until the first repeated number is found. Then it prints out in ascending order from 0 the table entries visited.

Make note of the values printed out and try some experiments to support the hypothesis:

Theorem 1 *In a table of size n and a skip of value k, the values printed out are $0, d, 2d, \ldots$ where d is the greatest common divisor of n and k.*

Try to prove this statement. Explain how this statment implies the previous statement about odd skips in tables of size 2^i, i any integer.