University of Washington

Department of Computer Science and Engineering

CSE 373

Problem Set 4 Solution

1. (4 points) Hash Table Insertion

```
0:                              0: 9679
1: 4371                         1: 4371
2:                              2: 1989
3: 1323 -> 6173                 3: 1323
4: 4344                         4: 6173
5:                              5: 4344
6:                              6:
7:                              7:
8:                              8:
9: 4199 -> 9679 -> 1989         9: 4199
  ( PART A )                      ( PART B )

0: 9679                         0:
1: 4371                         1: 4371
2:                              2:
3: 1323                         3: 1323
4: 6173                         4: 6173
5: 4344                         5: 9679
6:                              6:
7:                              7: 4344
8: 1989                         8:
9: 4199                         9: 4199
  ( PART C )                      ( PART D: 1989 can't be inserted )
```

2. (2 points) Memory Efficient Dictionary Approximation

   (a) True. If a cell is marked 0, then no word in the dictionary hashes to it.

   (b) False. If a cell is marked 1, then *some* item in the dictionary hashes to it. However, there is no guarantee that it is the word that we are looking up. For example, suppose we are using a very bad hash function that sends all words starting with "a" to position 1. Then position 1 would contain a 1, because the word "aardvark" hashes there. However, "agqk" would also hash to position 1 and hence be called correct by the algorithm.

3. (4 points) Hashing Strategy Disadvantages

   Open Hashing: Use of linked lists requires more storage overhead and more complex code.

   Linear Probing: Primary clustering.

   Quadratic Probing: Secondary clustering, and unless care is taken, it may be impossible to insert some items even when the table is not full.

   Double Hashing: Hash functions are usually fairly expensive to compute, so many simpler probes can be done in the time it takes to compute the secondary hash function.

4. (10 points) Hash Table Implementation

   The sample solution to this problem is stored separately on the course web site so it is easier to download and compile. See the bullet for Homework 4 on the "Homework" page.