1 Introduction

This lab will demonstrate the effect of the hash function on the performance of a hash table. You will examine the effect of one simple hash function in detail, then try to devise your own.

Using Eclipse, import “Lab11” in the following directory:

/afs/cs.uwm.edu/users/classes/cs351/401/pantherid/git/lab11.git

2 Hashing

A hash table stores values in a table (array) and uses a hash function to generate integral indexes from keys. The key is passed through the hash function which returns an integer. This integer modulo the table size is used as index which determines where a specific value will be stored. Because there are frequently more possible keys than spaces in the array, it is possible for two keys to hash to the same array element. This is called a collision. There are many ways to deal with collisions; in this lab we will use linear probing, searching forward in the table until an open space is found.

3 A Simple Hash Function

We will be storing Person objects in our hash table. A Person object has a first and last name and implements the methods hashCode() and equals() that will be used to determine the hash table array index in which to store the Person.

The current implementation of hashCode() for Person simply returns the ASCII value of the first character of its lastName field, which is set as the last name of its associated Person. The hash table will generate the initial array index in which to store the Person using the absolute value of the result of hashCode() modulo the capacity of the hash table (which is initially 11). If the initial array index produced is already occupied, subsequent array locations will be examined until an open location is seen in which to store the object.

In the driver supplied, the following values are put into the hash table. Next to them, write the hash values for the current implementation. You may use the table on below the list of entries.

Casey McKehee :

Corey Post :

Gregg Zaun :

Alcides Escobar :

Imad Zabalawi :

| K - 75 | L - 76 | M - 77 | N - 78 | O - 79 | P - 80 | Q - 81 | R - 82 | S - 83 | T - 84 |
| U - 85 | V - 86 | W - 87 | X - 88 | Y - 89 | Z - 90 | a - 97 | b - 98 | c - 99 | d - 100 |
| e - 101 | f - 102 | g - 103 | h - 104 | i - 105 | j - 106 | k - 107 | l - 108 | m - 109 | n - 110 |
| o - 111 | p - 112 | q - 113 | r - 114 | s - 115 | t - 116 | u - 117 | v - 118 | w - 119 | x - 120 |
Now draw the contents of the table as the Person objects are put in the table in the same order as above (see the put method in the Table class for further clarification).

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
</table>

What was the number of collisions? ____________________________________________________________
Run the driver and verify that you had filled out the contents correctly.

4 A better hash function

The number of collisions is directly related to the hashCode implementation of the key. Exactly which hashCode implementation is best depends on both the distribution of the data and size of the hash table. Change the implementation of hashCode so that it produces half the number of collisions (or fewer) as the existing implementation.

Casey McKeehe : ____________________________________________________________
Corey Post : ____________________________________________________________
Gregg Zaun : ____________________________________________________________
Alcides Escobar : ________________________________________________________
Imad Zabalawi : _________________________________________________________

How many collisions does it generate? ______________________________________________________

5 Eclipse hash function

Comment out the equals and hashCode methods from Person.java. Let Eclipse generate these methods by right clicking on the editor while Person.java is open and selecting Source > Generate hashCode() and equals(). ..... How many collisions are generated when this implementation is used?

When you are happy with your hashCode implementation, see your TA for the lab checkoff.