Computing Hash Functions using Python

Homework #9

By

Mario Pendleton

CS 303 [Algorithms and Data Structures](https://uab.instructure.com/courses/1507655)

October 27, 2019

### Problem Specification

. Implement a hashing function.

. Use the hashing function to create a hashmap implementation.

**Problem1.**

Create a hashmap that is made of elements HashElement(int Key, String Value). The size of hashmap will be 100. Implement the following methods for the hashmap:

a. put(int Key, String Value): Puts the key value pair in the hashmap at a certain index. You need to implement a simple hash function H(Key)=Key mod mapsize to find the index where you will put the pair. If collision occurs, i.e., a pair already exists in that index and the key is not the same as the current key, then you will use this function to resolve the collision, H(x)=(7\*H(x)+1) mod mapsize, until you get an empty slot. If the index is already full and the keys are the same, just replace the old value with the new one.

b. get(int Key): Gets the value associated with the Key. You should not do linear search throughout the hashmap for the key, rather you will calculate the index using the hash function stated above, go directly to that index and retrieve the value.

2. Write a driver program to test your implementation of hashmap. Allow the user to put or get data.

3. Implement linear probing to put a new value in your HashMap. The sequence of probes are: H(x) mod mapsize, (H(x) +1) mod mapsize, (H(x)+2) mod mapsize, (H(x)+3) mod mapsize … and so on.

4. Implement quadratic probing to put a new value in your HashMap. The sequence of probes are: H(x) mod mapsize, (H(x) +1) mod mapsize, (H(x)+4) mod mapsize, (H(x)+9) mod mapsize … and so on.

5. Use the hashmap implementation from the previous question (make sure to update the hashmap size) to store the given input file that consists of two fields: a UPC key and the corresponding description. Use the hashmap created to find the description associated with a given set of UPC keys. The input file UPC.csv provides the key and corresponding descriptions in a comma separated file and the various search keys are provided in the file input.dat. First test the program by entering couple of keys manually and print the description. Once you are convinced the program is working correctly, test the program for the given search keys and determine the total time taken to complete the search.

Compare the times for searching the keys using the given function, linear probing, quadratic probing with hashmap.

### Program Design

The following steps were required to develop this program:

1. Write a node class in python.

Class Node:

\_\_init\_\_(self, key, data)

Methods within the class.

printNode(self) – Prints a node stats.(key, data, left, right, parent)

1. Write a binary tree class in python.

Class HashMap

\_\_init\_\_(self)

size

map

Methods with the class.

put(self, key, data) – Inserts a node into a hash table

linear\_probe(self,key,data) – Linear search for next open index

linear\_probe(self,key,data) – Quadratic search for next open index

putq(self,key,data) – Inserts a node using a Quadratic Probe

get(self,key,data) – Returns the index of a given key and data give.

load(self,path) – Loads data from a file and places it into a hash tree.

loadCompare(self, path) – Reads a file and searches for values in a hash table.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Test Case | Input Values | Search Keys | Search Results | Reverse Bubble Sort |
| (f) | load(path) path = input.dat.txt  Table Size 10000 | 79,,INDIANA LOTTO  93,,treo 700w  123,,Wrsi Riversound cafe cd  161,,Dillons/Kroger Employee Coupon ($1.25 credit)  2140000070,,Rhinestone Watch  2140118461,,"""V"": Breakout/The Deception VHS Tape"  2144209103,VHS,Tintorera - Tiger Shark  2144622711,,Taxi : The Collector's Edition VHS  2147483647,,Toshiba 2805 DVD player  2158242769,288/1.12Z,GREEN SUGAR COOKIES4276  2158561631,,HOT COCOA W/BKMK  2158769549,njhjhn,gjfhjbgkj  2160500567,2.25 oz (64)g,Dollar Bar Rich Raspberry  2172307284,,Mixed seasonal flower bouquet  2177000074,,4 way 13 AMP Extension Lead (Wilkinson UK)  2184000098,21 oz,Christopher's Assorted Fruit Jellies  2187682888,,fairway | Process Time | Start: 0.04414 , End: 0.044152 | Total processing time: 1.1999999999998123e-05  ('Your entry', ['79', ' INDIANA LOTTO'], 'index:', 19)  Process Time | Start: 0.044196 , End: 0.044206 | Total processing time: 1.0000000000003062e-05  ('Your entry', ['93', ' treo 700w'], 'index:', 20)  Process Time | Start: 0.04424 , End: 0.04425 | Total processing time: 9.999999999996123e-06  ('Your entry', ['123', ' Wrsi Riversound cafe cd'], 'index:', 21)  Process Time | Start: 0.04428 , End: 0.044292 | Total processing time: 1.1999999999998123e-05  ('Your entry', ['161', ' Dillons/Kroger Employee Coupon ($1.25 credit)'], 'index:', 22)  Process Time | Start: 0.04432 , End: 0.044332 | Total processing time: 1.2000000000005062e-05  ('Your entry', ['2140000070', ' Rhinestone Watch'], 'index:', 23)  Process Time | Start: 0.044362 , End: 0.044372 | Total processing time: 1.0000000000003062e-05  ('Your entry', ['2140118461', ' "V": Breakout/The Deception VHS Tape'], 'index:', 24)  Process Time | Start: 0.044404 , End: 0.044416 | Total processing time: 1.1999999999998123e-05  ('Your entry', ['2144209103', 'VHS Tintorera - Tiger Shark'], 'index:', 25)  Process Time | Start: 0.044444 , End: 0.044456 | Total processing time: 1.2000000000005062e-05  ('Your entry', ['2144622711', " Taxi : The Collector's Edition VHS"], 'index:', 26)  Process Time | Start: 0.044486 , End: 0.044498 | Total processing time: 1.2000000000005062e-05  ('Your entry', ['2147483647', ' Toshiba 2805 DVD player'], 'index:', 27)  Process Time | Start: 0.044528 , End: 0.04454 | Total processing time: 1.2000000000005062e-05  ('Your entry', ['2158242769', '288/1.12Z GREEN SUGAR COOKIES4276'], 'index:', 28)  Process Time | Start: 0.044568 , End: 0.04458 | Total processing time: 1.1999999999998123e-05  ('Your entry', ['2158561631', ' HOT COCOA W/BKMK'], 'index:', 29)  Process Time | Start: 0.044608 , End: 0.044622 | Total processing time: 1.4000000000000123e-05  ('Your entry', ['2158769549', 'njhjhn gjfhjbgkj'], 'index:', 30)  Process Time | Start: 0.04465 , End: 0.044664 | Total processing time: 1.4000000000000123e-05  ('Your entry', ['2160500567', '2.25 oz (64)g Dollar Bar Rich Raspberry'], 'index:', 31)  Process Time | Start: 0.044694 , End: 0.044706 | Total processing time: 1.2000000000005062e-05  ('Your entry', ['2172307284', ' Mixed seasonal flower bouquet'], 'index:', 32)  Process Time | Start: 0.044736 , End: 0.04475 | Total processing time: 1.4000000000000123e-05  ('Your entry', ['2177000074', ' 4 way 13 AMP Extension Lead (Wilkinson UK)'], 'index:', 33)  Process Time | Start: 0.044778 , End: 0.044794 | Total processing time: 1.6000000000002124e-05  ('Your entry', ['2184000098', "21 oz Christopher's Assorted Fruit Jellies"], 'index:', 34)  Process Time | Start: 0.044832 , End: 0.044846 | Total processing time: 1.4000000000000123e-05  ('Your entry', ['2187682888', ' fairway'], 'index:', 35)  Process Time | Start: 0.042638 , End: 0.044892 | Total processing time: 0.002253999999999999 | 0.002253 |
| (f) | load(path) path1 = UPC.csv  Table Size 10,000 |  | Process Time | Start: 13.92676 , End: 13.92677 | Total processing time: 9.999999999621423e-06  ('Your entry', ['79', ' INDIANA LOTTO'], 'index:', 19)  Process Time | Start: 13.92685 , End: 13.926858 | Total processing time: 7.999999999341867e-06  ('Your entry', ['93', ' treo 700w'], 'index:', 20)  Process Time | Start: 13.926886 , End: 13.926894 | Total processing time: 8.000000001118224e-06  ('Your entry', ['123', ' Wrsi Riversound cafe cd'], 'index:', 21)  Process Time | Start: 13.926948 , End: 13.926956 | Total processing time: 8.000000001118224e-06  ('Your entry', ['161', ' Dillons/Kroger Employee Coupon ($1.25 credit)'], 'index:', 22) | 13.887 |
| (f) | load(path) path2 = UPC\_unsorted.cvs  Table Size 10,000 |  | None Found – Table must be enlarged | 13.892 |

1. Use the a method to read the following txt files and covert them into arrays to be sorted

* input\_100.txt
* UPC.csv
* UPC\_unsorted.csv
* Input.dat.txt

### Testing Plan

Sample nodes were created and printed. New instance of hash map was created to house these nodes. The nodes were then placed into the new hash map with both string and integerkeys. Once the nodes were in the map I tested the get() function. I then tested the search node function for nodes that did not exist and nodes that were improperly keyed. Once the core functions were working I created 2 import functions. One function loads a file into a tree and the other loads and file and compare the values to the values in the tree.

### Test Cases

The test cases are shown in the table below using a MacBook Pro 16GB, 8 core 2.3 GHz Intel Core i9:

### Analysis and Conclusions

### My analysis showed no difference in speed between the quadratic probe and linear. I only found that the quadratic probe continuously fell outside the bounds of the table.

### References

Textbook, python.org, and examples provided in the assignment.

**Screen Shots**

****

