BLG335E, Analysis of Algorithms 1

Name:	İbrahim KARAHAN
Student Number:	150160550
CRN:	10983
Project 3	

1. Introduction

In this homework, I try to implement several hash function according to given input file which name is vocab.txt. According to this input file key and value pairs I create my hash table and place the value to the hash table with using the hash function output. When output of hash function is same as the before word in the vocab.txt then we try to solve this situation which name is collusion. In collusion problem I used an open addressing method. Open addressing method is linear probing, double hash and universal hash.

In my code base, I create two class. One of them is vocab.txt for keeping value and the other necessary things. The other one is for serarch.txt. I create three array. First one for keeping the vocab object and size is 13423. Second one for keeping the search object and size is 13423. The last one for hash table and size is 100K. I choose 100K because this size is an optimum. I find the size manually. In my code, I give some comment and make clear my code and also try to explain some part of the misunderstanding.

2. Result

Result of linear hashing with m size is 17863.

Figure-1

Result of linear hashing with m size is 21929.

Figure-2

Result of double hashing with m size is 17863.

Figure-3

Result of double hashing with m size is 21929.

Figure-4

Result of universal hashing with m size is 17863.

Figure-5

Result of universal hashing with m size is 21929.

Figure-6

2.1   Tables

Collusion Number for Hashing

	Linear Hash	Double Hash	Universal Hash
M=17863	0	0	5934649
M=21929	0	0	4034051

Table-1

Collusion Number for Searching

	Linear Search	Double Search	Universal Search
M=17863	169993388	169993388	461754009
M=21929	169993388	169993388	292819907

Table-2

3. Conclusion

To sum up, I learn the implementation and compilation of hashing algorithm. I observe the problem of collusion occurrence and learn how collusion problem is handled.

4. Compilation

When you debug my homework on ssh server, you should write below command line.

g++ Source.cpp -o output.o –std=c++11
./output.o vocab.txt search.txt

```
[karahanil6@ssh ~]$ g++ Source.cpp -o output.o -std=c++11
[karahanil6@ssh ~]$ ./output.o vocab.txt search.txt
Hash Number(M):17863
Linear Hashing & Searching
Collusion Number For Linear Hashing:0
this word held, is not in the vocab.txt
Collusion Number For Linear Hash Searching:169993388
*************************************
Double Hashing & Searching
Collusion Number For Double Hashing:0
this word held, is not in the vocab.txt
Collusion Number For Double Hash Searching:169993388
*************************************
 Universal Hashing & Searching
Collusion Number For Universal Hashing:6587162
this word recover is not in the vocab.txt
Collusion Number For Universal Hash Searching:240349525


Hash Number(M):21929
Linear Hashing & Searching
Collusion Number For Linear Hashing:0
this word held, is not in the vocab.txt
Collusion Number For Linear Hash Searching:169993388
*************************************
Double Hashing & Searching
Collusion Number For Double Hashing:0
this word held, is not in the vocab.txt
Collusion Number For Double Hash Searching:169993388
*************************************
Universal Hashing & Searching
Collusion Number For Universal Hashing:3849094
this word recover is not in the vocab.txt
Collusion Number For Universal Hash Searching:625668692
```

Figure-7