Eddie C. Fox

August 9, 2016

CS261

1. Give an example of two words that would hash to the same value using hashFunction1 but would not using hashFunction 2.

The two words “life” and “file”. Both have the same letters, but the difference in ordering will create a difference between the two hash functions.

2. Why does the above observation make hashFunction2 superior to hashFunction1?

Unlike hashFunction1, hashFunction2 cares about the ordering of the letters. Because of this, there will be more unique hash indexes generated which will reduce hash collision. Buckets will therefore have less link chains overall.

3. When you run your program on the same input file once with hashFunction1 and once with hashFunction2, is it possible for your hashMapSize function to return different values?

No, both hash functions create hash maps with the same size.

4. When you run your program on the same input file once with hashFunction1 and once with hashFunction2, is it possible for your hashMapTableLoad function to return different values?

No. While it is possible that hashFunction1 and hashFunction2 might have different distributions of links, that would not affect the number of buckets and the number of links in the hash map, which are the only two relevant factors when calculating the load factor of the hash map.

5. When you run your program on the same input file once with hashFunction1 and once with hashFunction2, is it possible for your hashMapEmptyBuckets function to return different values?

Yes. Because the hash functions will likely distribute the links differently, there is a possibility that the number of empty buckets will be different and that the buckets which are empty will be different.

6. Is there any difference in the number of empty buckets when you change the table size from an even number like 1000 to a prime like 997?

Changing the table size from an even number to a prime number seems to slightly increase the number of empty buckets. This would indicate more efficient performance, because if there are more buckets compared to the same number of links, the load factor will be slightly lower, leading to slightly better performance.