In-class Exercise Week 5

Questions

1. Demonstrate what happens when we insert the keys 5, 28, 19, 15, 20, 33, 12, 17, 10 into a hash table with collisions resolved by chaining. Let the table have 9 slots, and let the hash function be $h(k) = k \ mod \ 9$.

2. Let H be a hash table where collisions are handled by chaining. Re-hashing is used each time the load factor, α exceeds 0.5. Assume that the initial size of H is 2 and that re-hashing doubles the size of the table. After inserting 10 items with different keys, what is the size of the hash-table H?

3. a) Suppose we design an algorithm for hashing character based item, such as strings. Hence word '**cat**' i.e $99 + 97 + 116 = 312$ can be thought of as sum of ordinal values . Write suitable code/pseudocode which takes string and table size and returns hash value in the range 0 to table size-1.

   b) Would there be collision for anagrams? Propose a solution to handle such cases. Anagram is word or phrase formed by rearranging the letters of a different word or phrase, typically using all the original letters exactly once.

   For example *"rail safety" = "fairy tales".*

Solutions

1. Demonstrate what happens when we insert the keys 5, 28, 19, 15, 20, 33, 12, 17, 10 into a hash table with collisions resolved by chaining. Let the table have 9 slots, and let the hash function be $h(k) = k\ mod\ 9$.

Solution:

Label the slots of the table to be 0, 1, 2, …, 8. Numbers which appear to the left in the table have been inserted layer.

0	
1	10 19 28
2	20
3	12
4	
5	5
6	33 15
7	
8	17

Solutions

2. Let H be a hash table where collisions are handled by chaining. Re-hashing is used each time the load factor, α exceeds 0.5. Assume that the initial size of H is 2 and that re-hashing doubles the size of the table. After inserting 10 items with different keys, what is the size of the hash-table H?

Solution:

α after 1st insert is $\frac{1}{2}$.

α after 2nd insert is $\frac{1+1}{2*2} = \frac{2}{4}$.

α after 3rd insert is $\frac{2+1}{4*2} = \frac{3}{8}$.

α after 4th insert is $\frac{3+1}{8} = \frac{4}{8}$.

Continuing in this manner …

α after 10th insert is $\frac{10}{32}$.

Since α is the number of items in the table divided by the size of the table, the size of the table after 10 insertions is 32.

Solutions

3. a) Suppose we design an algorithm for hashing character based item, such as strings. Hence word '**cat**' i.e $99 + 97 + 116 = 312$ can be thought of as sum of ordinal values . Write suitable code/pseudocode which takes string and table size and returns hash value in the range 0 to table size-1.

   b) Would there be collision for anagrams? Propose a solution to handle such cases. Anagram is word or phrase formed by rearranging the letters of a different word or phrase, typically using all the original letters exactly once.

   For example *"rail safety" = "fairy tales".*

Solution:

a)

```
def hash(astring, tablesize):
    sum=0
    for pos in range(len(astring)):
        sum=sum + ord (astring[pos])
    return sum%tablesize
```

b) Take weights (position of string) too as a factor. Hence sum can be written as

$1 * 99 + 2 * 97 + 3 * 116 = 641$