Hashing

Adapted By Manik Hosen

Basic Terminology

- Question: Define Hashing.
- Ans: Concept of building a data structure that can be searched in O(l) time is called Hashing.
- Question: Define Hash Table with example.
- Ans: A Hash Table is a collection of items which are stored in such a way as to make it easy to find them later. This is a hash table with 4 slots:

0	1	2	3
None	None	None	None

- Question: Define Slot.
- Ans: Each position of the hash table is called a slot. A slot can hold an item and is named by an integer value starting at 0.

Hash Function

- Question: Define hash function and discuss.
- Ans: The mapping between an item and the slot where that item belongs in the hash table is called the Hash Function.
- The has function will take any item from the collection and return an integer in the range if slot names between 0 to m-1. Assume that we have the set of integer items 54, 26, 93, 17, 77, 31. In 'reminder method' simply taken an item and divided it by m and the hash function return hash value (h(item)=item%m). Where m is the size of table.

Hash Function

- Question: Find hash values for given items below of a hash table of size 6:

54	26	93	17	31

- Ans: h(54)=54%6=0 h(26)=26%6=2
  h(93)=93%6=3 h(17)=17%6=5
  h(31)=31%6=1

Hence, The Hash values are:

Hash Value	0	1	2	3	4	5
Item	54	31	26	93	None	17

Basic Terminology

- Question: Define Load Factor.
- Ans: In a Hash Table ratio of occupied slots and table size is called LoadFactor. It is denoted by λ.
- Question: Define Collision.
- Ans: If we find that we need to insert two or more different items in the same slot then that is called Collision.

Hash Function

- Question: Define Perfect Hash Function.
- Ans: A hash function that maps each item into a unique slot is called a Perfect Hash Function.
- Question: Discuss Folding Method for constructing hash function.
- Ans: The folding method for constructing hash function begins by dividing the item into equal size pieces. Then they are added together and being divided by the total number of slots in the hash table.
- For example: If we consider a phone number 436-555-4601 it will be divided as 43,65,55,46,01. Then adding these we get 210. If size of hash table be 11 then 210%11=1. Hence it will be in slot 1.

Hash Function

- Question: Discuss Mid-Square Method for constructing hash function.
- Ans: In Mid-Square Method we first square the item and then extract a portion of the resulting number and finally perform the reminder step.. For Example: If the item is 44. First we calculate 44^2=1936. Then we will take a part of 1936. If we extract middle two digit we find 93. After performing reminder step we get 93%11=5. This will be the hash number.
- Question: Can we creat Hash Function for string?
- Ans: Yes, we can.

Hash Function

- Question: Find Hash Value of the following Items for table size 11 using Mid-Square Method.

54	26	93	17	77	31

- Ans: $54^2=2916$ $26^2=676$ $93^2=8649$

  $17^2=289$ $77^2=5929$ $31^2=961$

If we take the mid digits(except one in both side) we get 91, 7, 64, 8, 92, 6.

After reminder step we get 3, 7, 9, 8, 4, 6. Hence the Hash Values are 3, 7, 9, 8, 4, 6 and Hash Table is:

Hash Value	0	1	2	3	4	5	6	7	8	9	10
Item	-	-	-	54	77	-	31	26	17	93	-

Collision Resolution

- Question: Define Collision Resolution.
- Ans: Sometime two items hash to the same slot, there the process of placing the second item in the hash table is called Collision Resolusion.
- Question: Define Linear Probing.
- Ans: In open addressing collision resolution process we perform an open addressing technique. This is called Linear Probing.

Collision Resolution

- Question: What is Open addressing Process? Discuss.
- Ans: Open Addressing Process is a method for collision resolution. In this method first we find what item is making problem. After finding that we start search from the slot it should be placed. If we find any free slot then we place the item there.
- Question: What is the disadvantage of Open Addressing Process?
- Ans: A disadvantage of open addressing is the tendency of clustering. This means if many collisions occur at the same hash value, a number of surrounding slots will be filled by the linear probing resolution. This will have an impact on other items that are being inserted.

Collision Resolution

- Question: What is Plus 3 Probe? Discuss about it.
- Ans: One way to deal with clustering we skip slots, thereby more evenly distributing the items that have caused collisions. This is potentially reduce the clustering. Tis process is called Plus 3 Prob. This means that once a collision occurs, we will look at every third slot until we find one that is empty.

Collision Resolution

- Question: What is Rehashing?
- Ans: The process of looking another slot skipping one after collision is called Rehashing.
- Question: What is Quadratic probing? Discuss
- Ans: A variation of the linear probing idea is called quadratic probing. In this process ionstead of using a constant “skip” value, we use rehash function that increments the hash value by 1,3,4,7,9 and so on. This means that if the first hash become h then the successive are h+1, h+4, h+9 and so on. In other words quadratic probing uses a skip consisting of successive perfect squares.

Collision Resolution

- Question: What is Chaining? Discuss it.
- Ans: Another method for collision resolution allows each slot to hold a reference to a collection of Items. This is called Chaining. Chaining allows many items to exist at the same location in the hash table. When collision happen, the item takes placed in the proper slot in chain. As more and more the item hash to the same slot the difficulty of the collection increases.