TUTORIAL-5

HASHING:

Hashing is a technique of storing values and searching for them from the tables.

Each value is stored associated with a key.

Hashing permits us to find elements much faster.

This data structure organizes the elements in an order that is convenient for its own purpose.

Hashing computes an integer value for each object which is called the hash code.

To compute correct array index using the hash function is called “hashing the key to an array index”.

The hash function must be chosen so that its return value is always a valid index for the array.

Time complexity for insert, delete, search, update is O(1).

COLLISION:

If two distinct keys have got the same hash code then collision occurs because two table entries can not be stored at the same table address.
For example considering the hash function h(k) = k/10, hash code for key 23 is 2 [h(23) = 23/10 =2] and hash code for key 28 is also 2 [h(28)=28/10=2].
In this circumstance both the key will map to the same table address creating a collision.

COLLISION RESOLUTION METHOD:

1. Open addressing with linear probing:
2. Open addressing with double hashing
3. External chaining
4. Hash bucket

LOAD FACTOR:

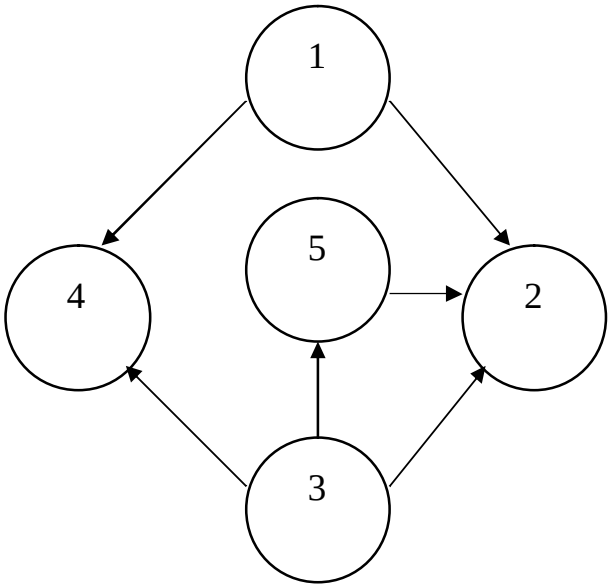
The load factor is related to the percentage of entries in the table that are occupied.
Load Factor=M/N where M = no of occupied entries and N = table size.
For example a table with a load factor of 0.30 means that the table is 30% full and a table with a load factor 0.75 means that the table is 75% full.

What is the drawback of linear probing?

Clustering is the drawback of linear probing. A cluster is a sequence of adjacent occupied entries in a hash table. Clustering is the tendency for keys to be clustered about an area in the list where one or more collisions have occurred. This problem can be solved using double hashing method.

GRAPH

Example-1:



Considering the above digraph:

a) What is the in-degree sequence?
In-degree sequence: 0,3,1,1,1 (number of arcs in)

b) What is the out-degree sequence?
Out-degree sequence:2,0,2,1,1 (number of arcs out)

c) What is the degree sequence?
Degree dequence: 2,3,3,2,2 (sum of in-degree and out-degree)

d) What are the set vertices adjacent from vertex 2?
There are no vertices adjacent from vertex 2

e) What are the set of edges incident from vertex 3?
(3,5) and (3,2) are the only edges incident from vertex 3.

f) What are the directed cycles and their lengths?

The above digraph has no directed cycles.

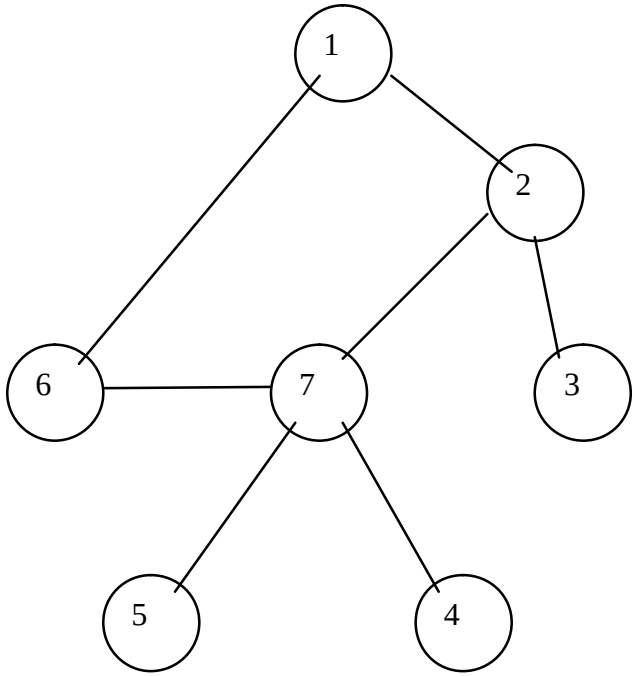
g) Is the digraph strongly connected?
No. because it is not possible to get from one vertex to any other vertex (eg. from vertex 2 to 1, 2 to 5, 2 to 3, 4 to 1 etc)

h) What is the diameter of the digraph?
The diameter of the digraph is undefined (8) because it is not strongly connected.

i) What is the distance from vertex 3 to 2?
The distance from vertex 3 to 2 is 1 because to calculate distance we shall consider the shortest path. (So in this case we shall not consider the path from 3 to 5 and 5 to 2)

j) What is the distance from vertex 1 to 3?
The distance from vertex 1 to 3 is undefined because it is not possible to get from vertex 1 to 3.

Example-2:



a) 6, 1, 2, 7, 4 is a walk ?
yes, it is a walk, because (6,1), (1,2), (2,7), (7,4) are edges of the graph.

b) 6, 1, 2, 7, 4 is a path ?
Yes, it is a path because no vertex is repeated.

c) 1, 2, 3, 4 is a path?
No, it is not a path in the graph because (3,4) is not an edge of the graph.

d) 7, 2, 3, 2, 1, 6 is a path?
Even though it is a walk in the graph, it is not a path because vertex 2 is repeated.

e) Is the graph connected?
Yes, it is a connected graph because it is possible to get from any vertex to any other vertex.

f) What is the diameter of the graph?
The diameter of the graph is 3 because we can get from one vertex to any other vertex by a path of at most 3 edges.

g) Is there any cycle in this graph?
Yes, there is a cycle in this graph which is 1, 2, 7, 6, 1