



**JAIN**  
DEEMED-TO-BE UNIVERSITY

SCHOOL OF  
COMPUTER  
SCIENCE AND IT

Master of Computer Applications

**Data Structures**

Module 5

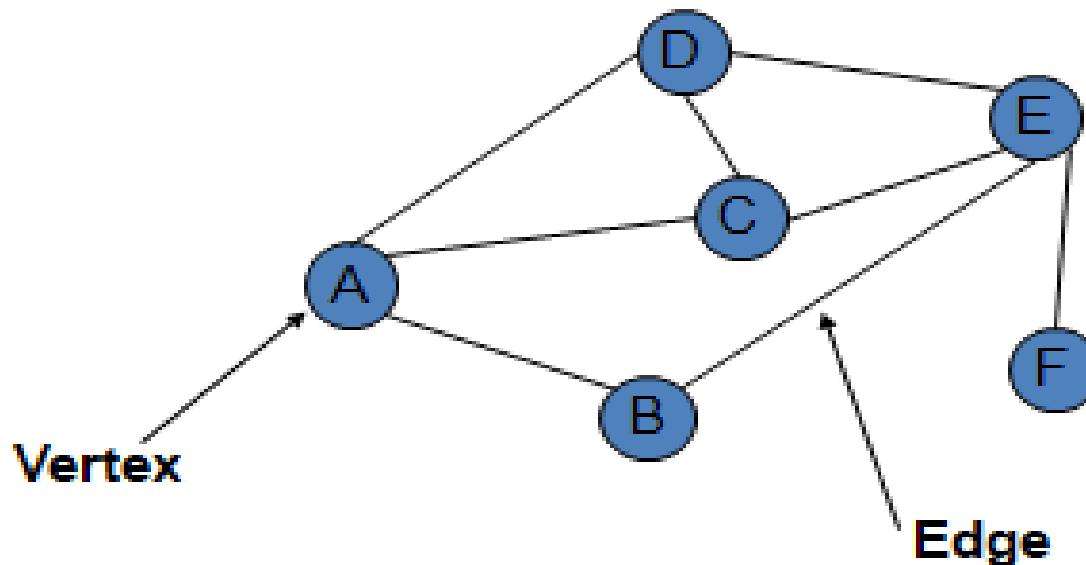
**GRAPH DATA STRUCTURES**

## Syllabus Contents – Graphs ADT

- Introduction to graph representation and Terminology
- Types of Graphs
- Graph traversal using Stack and Queue
- Applications of Depth First and Breadth First Traversal
- Applications of graph,
- Detect Cycle in a Directed Graph and in an undirected graph,
- Transitive Closure of a Graph using DFS.
- Topological sorting of Directed Acyclic Graphs.

- Consist of:
  - Vertices
  - Edges
  - it is an ordered pair of sets  $G(V,E)$  is called Graph.

Extremely useful tool in modeling problems



**Vertices** can be considered “sites” or locations.

**Edges** represent connections.

# Types of Graphs

- Directed Graphs

- Where there **is direction arrow** at the **end of the edge**

- Undirected Graphs

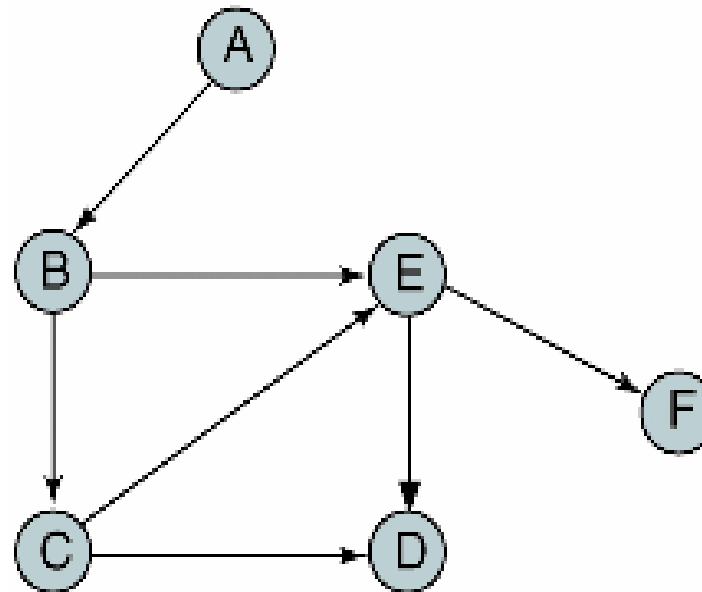
- **No arrows in edges** (Bi-direction between two nodes)

- Weighted Graphs

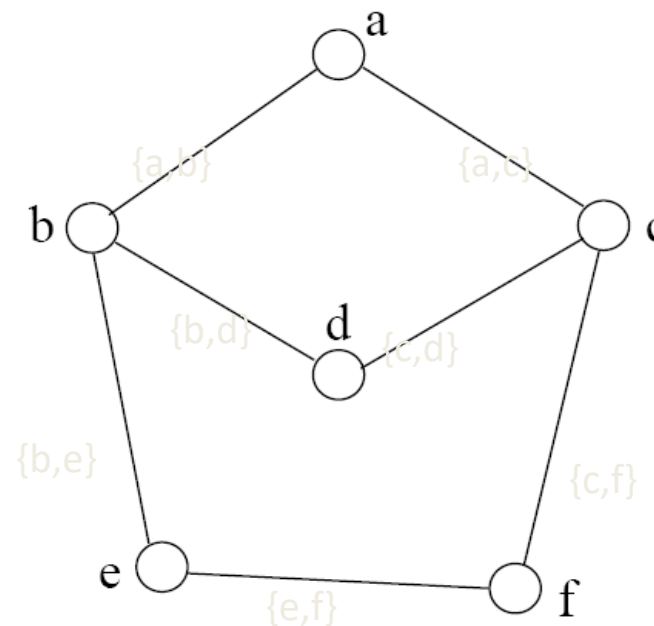
- **Cost will be assigned** in each edge in the graph

# Directed Graph

- A graph is directed if **direction is assigned to each edge**. We call the directed edges ***arcs***.
  - An edge is denoted as an ordered pair  $(u, v)$



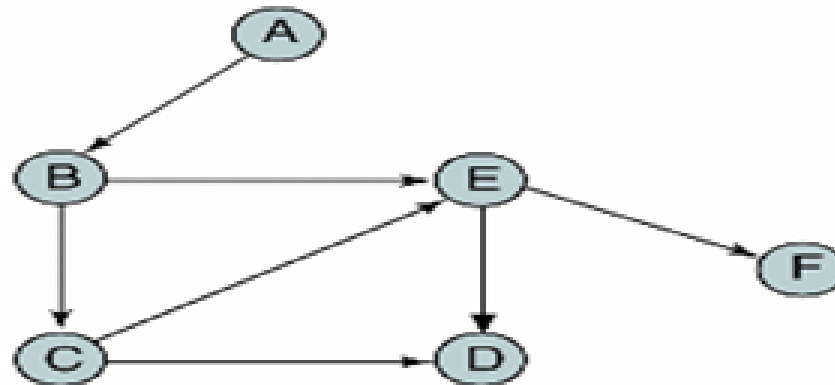
An undirected graph is **specified by an ordered pair (V,E)**, where V is the **set of vertices** and E is the **set of edges**



$$V = \{a, b, c, d, e, f\}$$

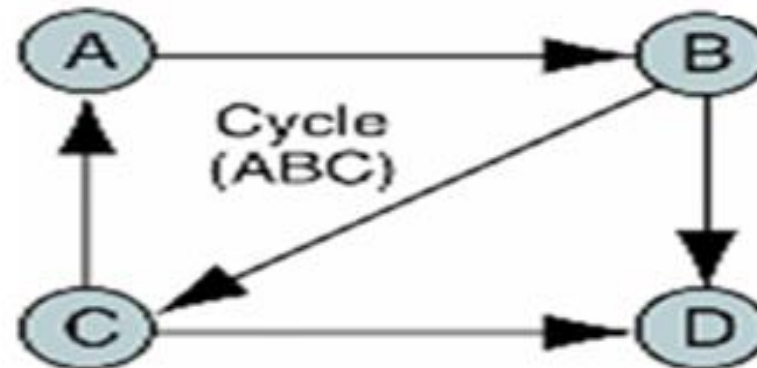
$$E = \{\{a, b\}, \{a, c\}, \{b, d\}, \{c, d\}, \{b, e\}, \{c, f\}, \{e, f\}\}$$

## 1. **Acyclic**: No circuit format of the Graph.

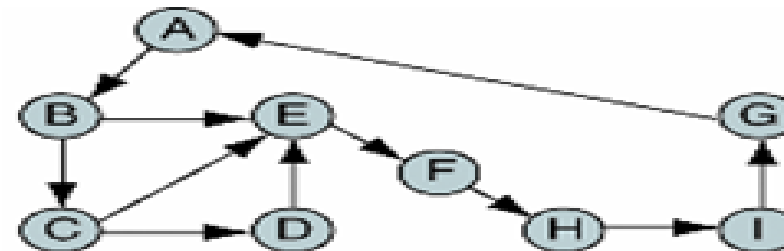


(a) **No cycle is found**

## 2. **Cycle**: A cycle is a path along the directed edges from a vertex to itself.

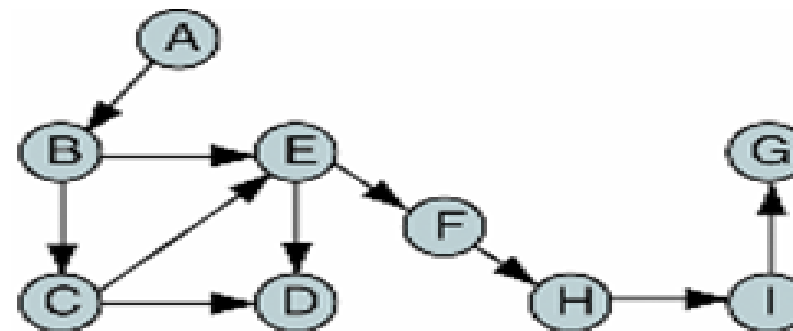


**3. Strongly Connected Graph:** It is strongly connected if there is a path from each vertex to every other vertex, considering direction.



(b) Strongly connected

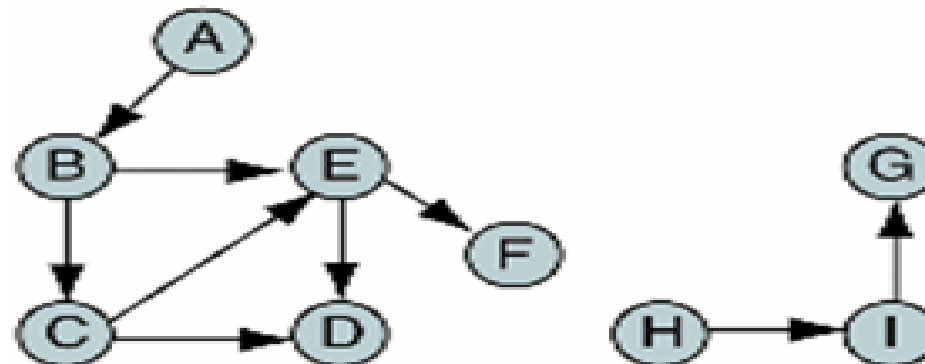
**4. Weakly Connected Graph:** It is strongly connected if there is a path from each vertex to every other vertex, considering direction, otherwise, it is weakly connected.



(a) Weakly connected

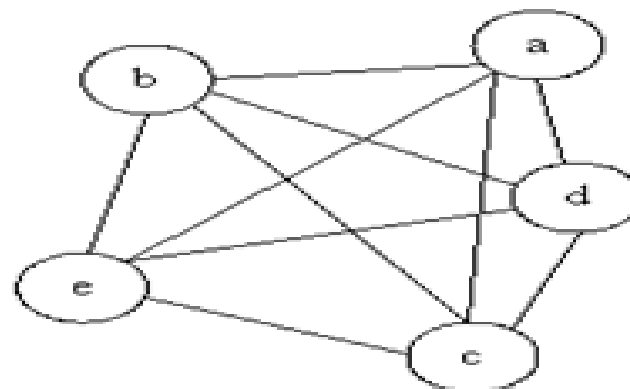


**5. Disjoint Graph:** A graph is disjoint if it is not connected



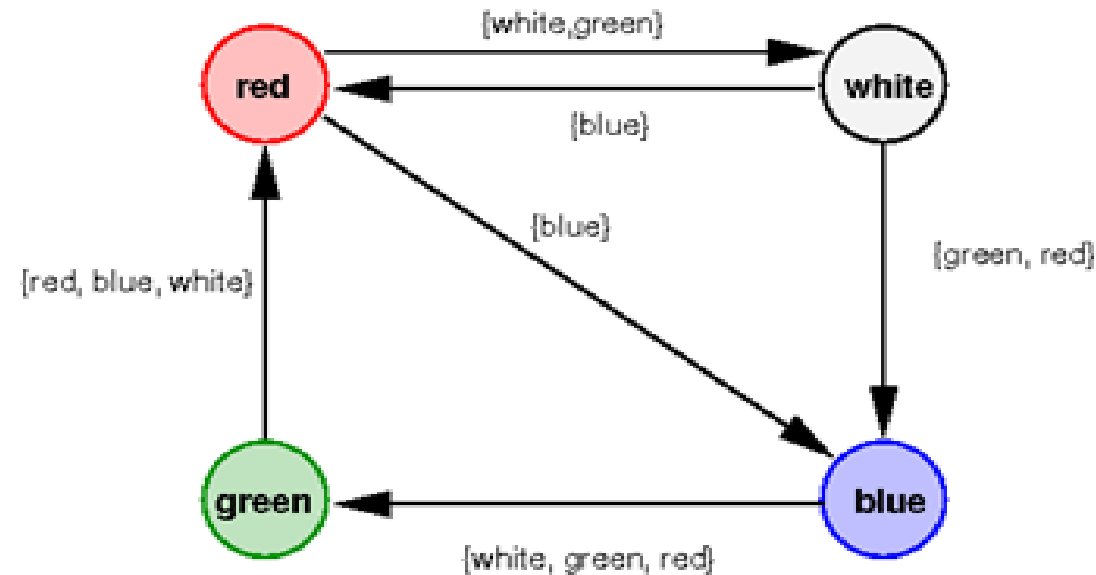
(c) Disjoint graph

**6. Complete Graph:** a graph that has the maximum number of edges, A graph in which every vertex is directly connected to every other vertex.

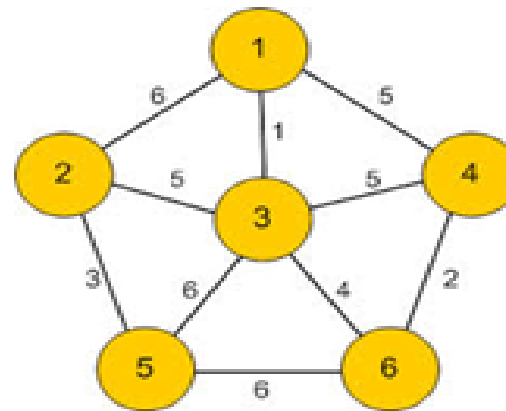


(e)

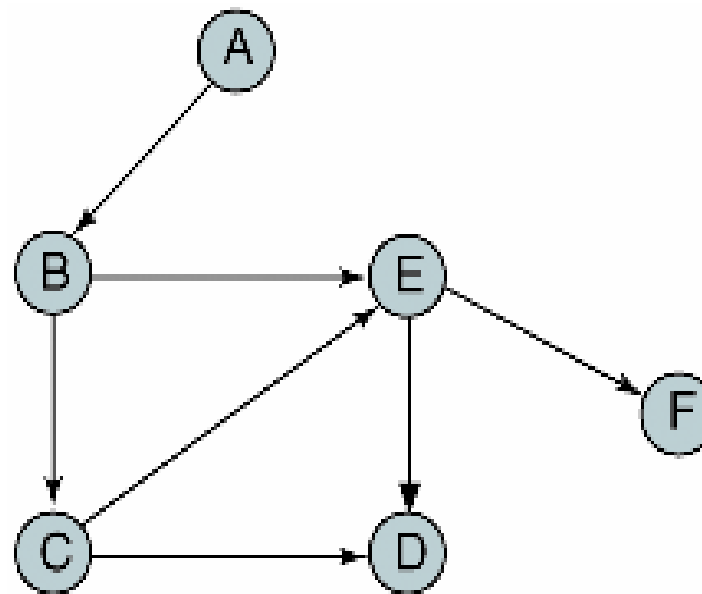
## 7. Labeled Graph: The Graph having names of the edges



## 8. Weighted Graph: The Graph having weights of the edges, A graph in which each edge carries a value.



**1. Path:** A sequence of vertices that **connects two nodes** in a graph.



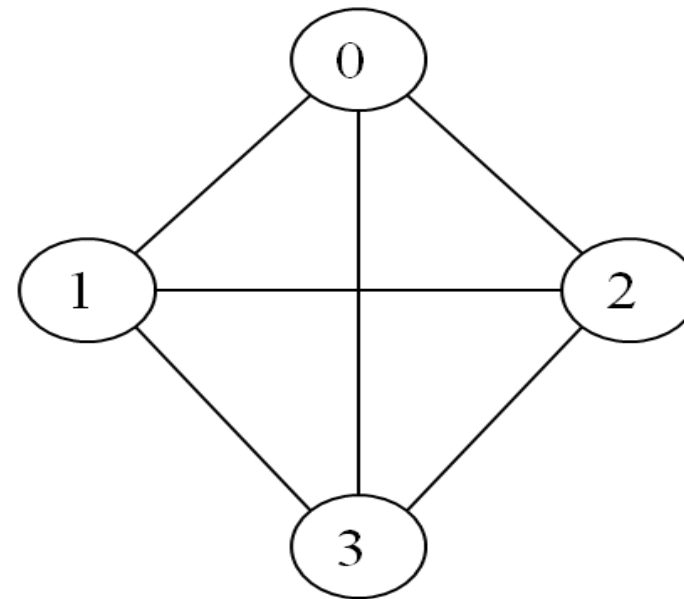
**Path**

**A->B->C->E->F,  
A->B->E->F**

## 2. Adjacent Vertices:

**Two vertices in a graph that are connected by an edge.**  
Sequence of vertices in which **each vertex is adjacent to next one**. Two vertices are **adjacent** (or **neighbors**) if there is a **direct path** connecting them

- ✓ 1 is adjacent to 0
- ✓ 1 is adjacent to 3
- ✓ 1 is adjacent to 2



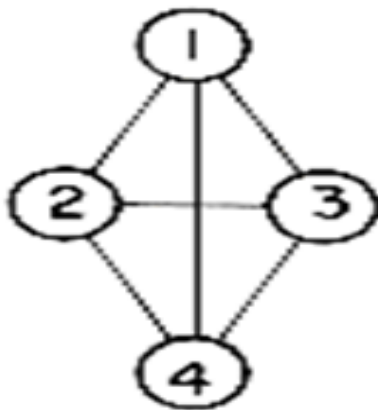
**G1**

$$V(G1) = \{0, 1, 2, 3\}$$

$$E(G1) = \{(0, 1), (0, 2), (0, 3), (1, 2), (1, 3), (2, 3)\}$$

### 3) Subgraph: $G'(V', E')$ is subgraph of $G(V, E)$

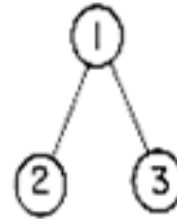
- $V(G') \subseteq V(G)$  and  $E(G') \subseteq E(G)$



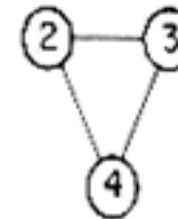
G1



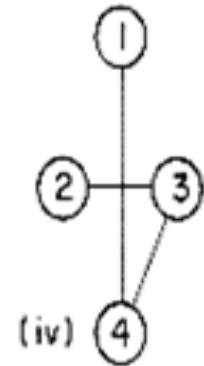
(i)



(ii)



(iii)

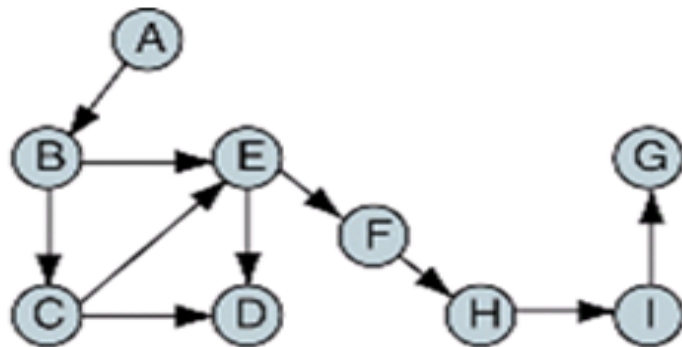


(iv)

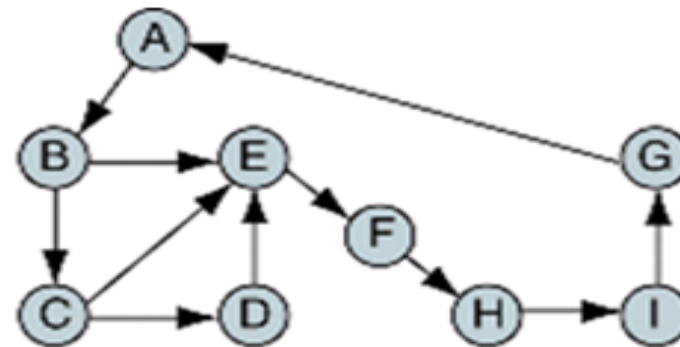
subgraphs of G1

**4) Degree of Graph:** The degree of a vertex is the number of lines incident to it.

- In Figure (a), the degree of vertex B is 3 and the degree of vertex E is 4



(a)

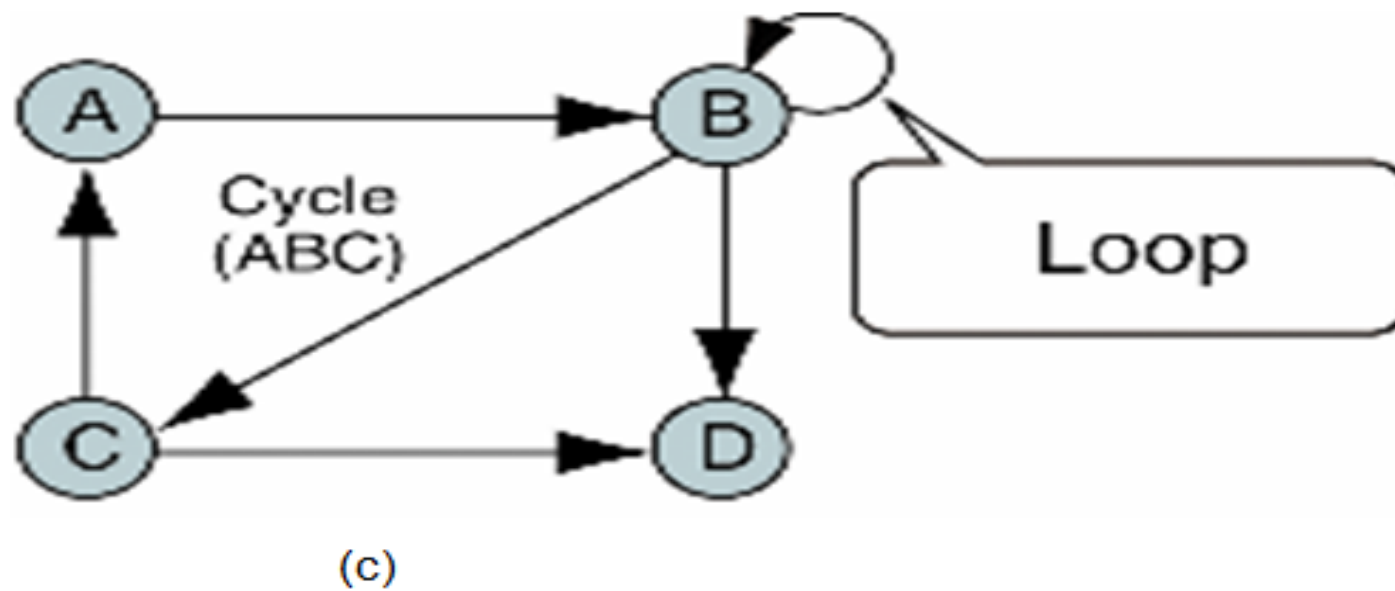


(b)

- The **outdegree** and **indegree** of a vertex
  - In Figure (a), the indegree of vertex B is 1 and its outdegree is 2
  - In Figure (b), the indegree of vertex E is 3 and its outdegree is 1

Activ

5. **Loop:** single arc begins and ends at the same vertex.



ABC is cycle

- Two **popular computer representations techniques** are available, but both represent the **vertex set and the edge set** in different ways.

## 1. Adjacency Matrix

Use a **2D matrix** to represent the graph

## 2. Adjacency List

Use a **1D array of linked lists**

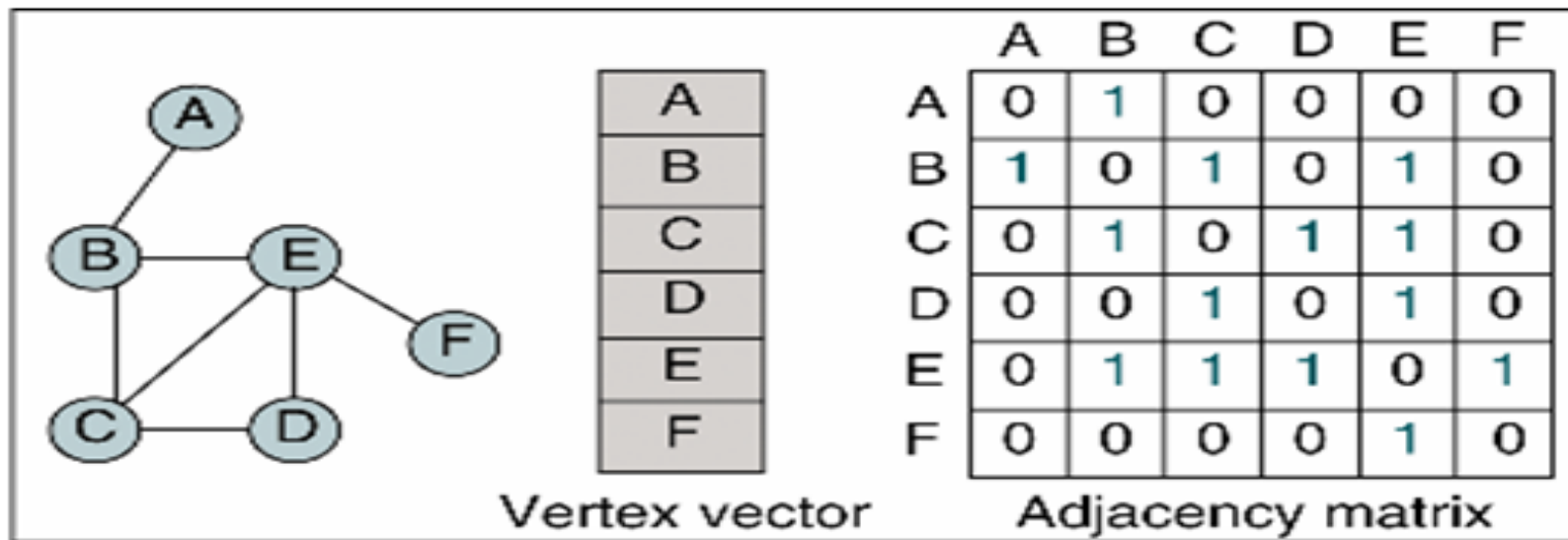


# Adjacency Matrix

Module No.5  
Non-Linear DS

## Adjacency Matrix with UnDirected Graph:

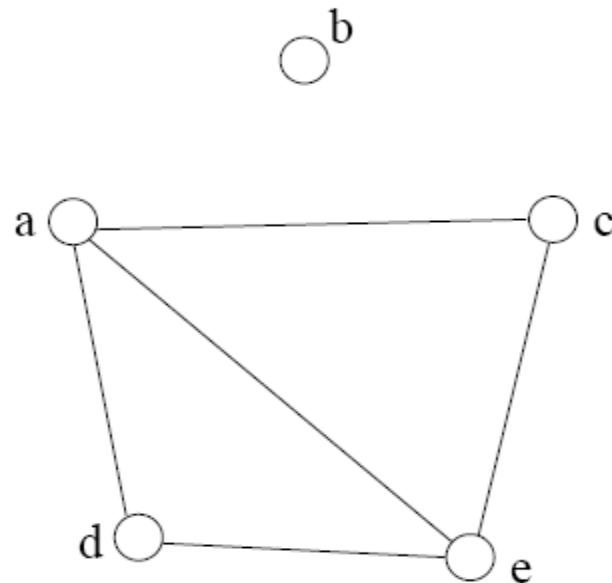
A two-dimensional matrix, in which the rows represent source vertices and columns represent destination vertices. Data on edges and vertices must be stored externally. Only the cost for one edge can be stored between each pair of vertices.



**(a) Adjacency matrix for nondirected graph**

Activ

# Adjacency Matrix

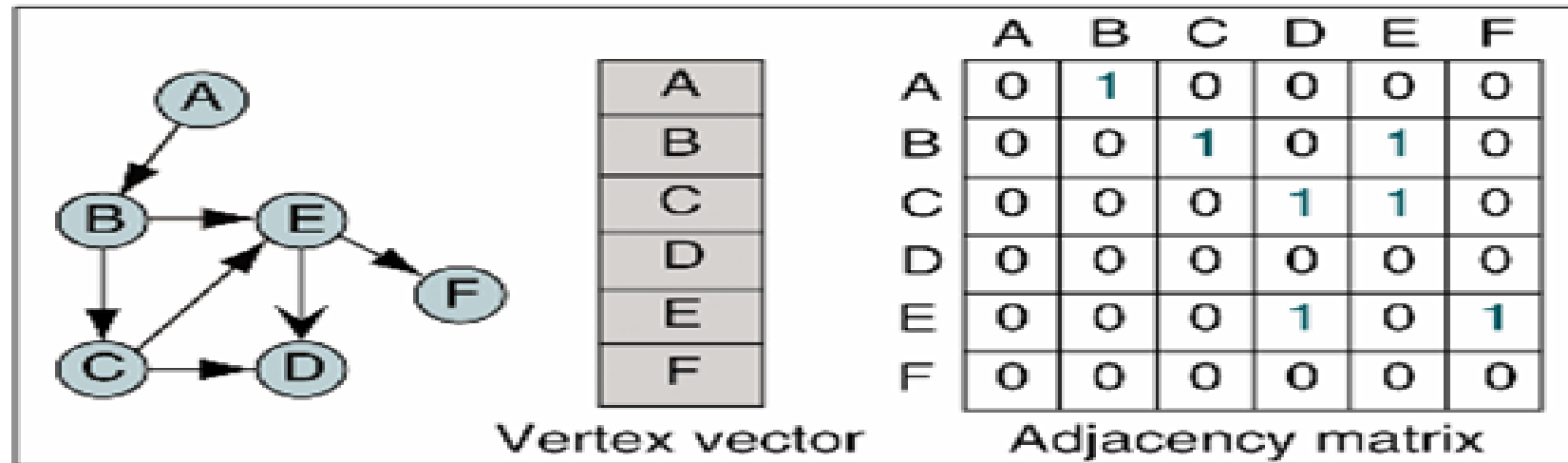


	a	b	c	d	e
a	0	0	1	1	1
b	0	0	0	0	0
c	1	0	0	0	1
d	1	0	0	0	1
e	1	0	1	1	0

- 2D array  $A[0..n-1, 0..n-1]$ , where  **$n$  is the number of vertices** in the graph
- Each **row and column** is indexed by the **vertex id**.
  - e.g  $a=0, b=1, c=2, d=3, e=4$
- An **array entry  $A[i][j]$  is equal to 1** if there is an edge connecting vertices  $i$  and  $j$ . Otherwise,  **$A[i][j]$  is 0.**
- The storage requirement is  **$\Theta(n^2)$ . Not efficient if the graph has few edges.**
- We can **detect in  $O(1)$  time** whether **two vertices are connected**.

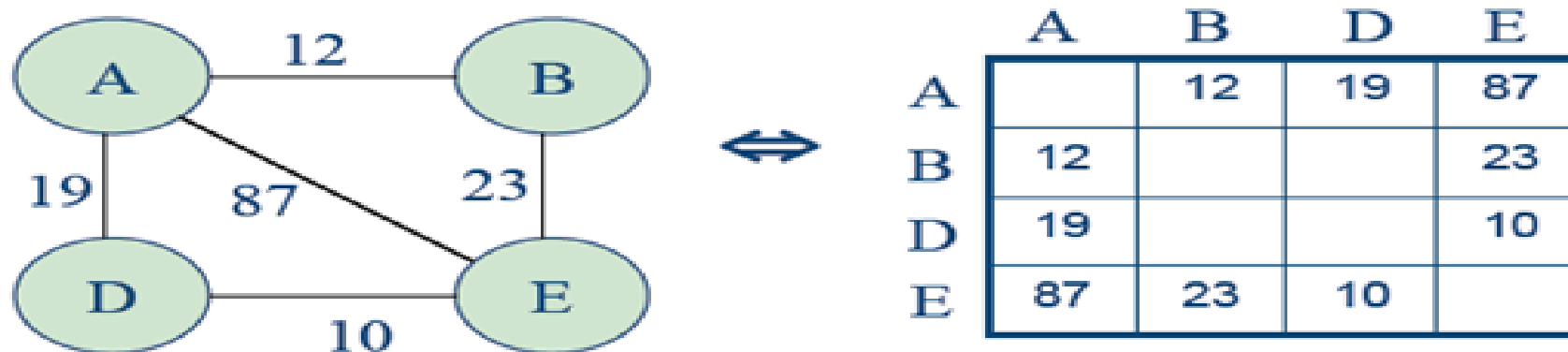
# Adjacency Matrix

Module No.5  
Non-Linear DS



**(b) Adjacency matrix for directed graph**

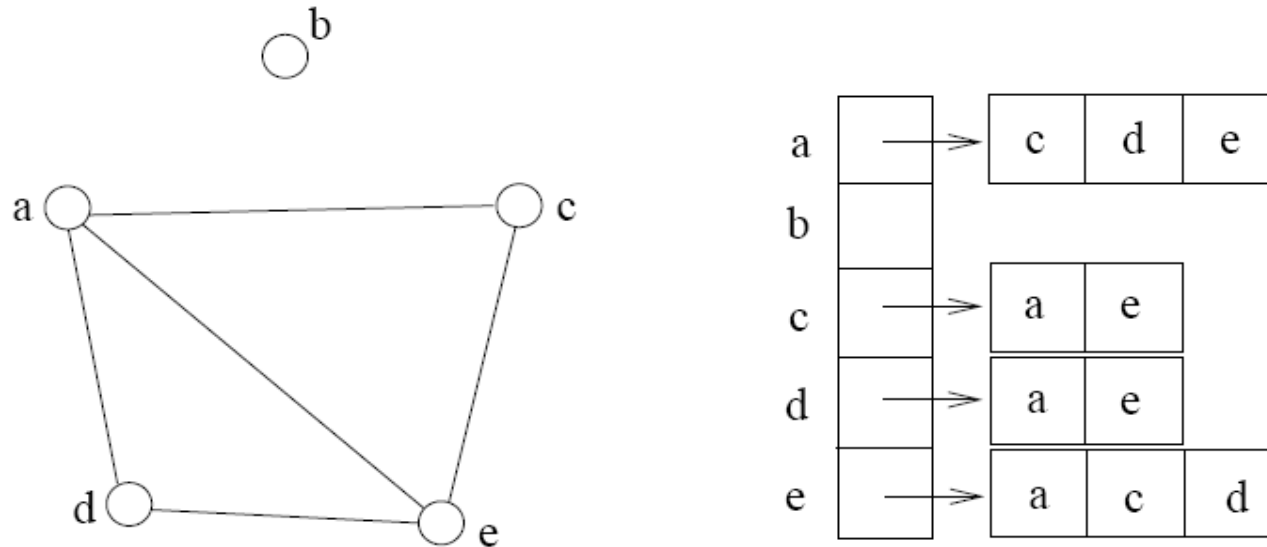
## ✓ c) Adjacency Matrix with Weighted Graph



Notice anything about this matrix?  
What would be the effect if the arcs were directed?

# Adjacency list

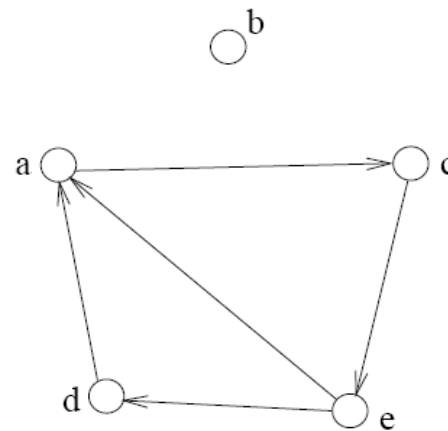
Module No.5  
Non-Linear DS



- The adjacency list is an array  $A[0..n-1]$  of lists, where  $n$  is the number of vertices in the graph.
- Each array entry is indexed by the vertex id (as with adjacency matrix)
- The list  $A[i]$  stores the ids of the vertices adjacent to  $i$ .

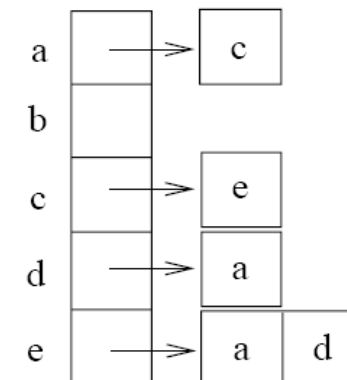
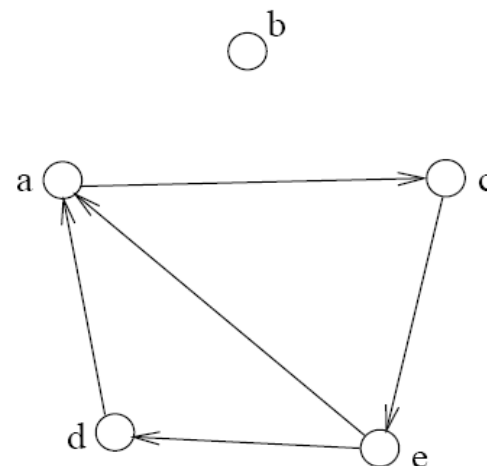
- The adjacency matrix and adjacency list can be used

## 1. Adjacency Matrix



	a	b	c	d	e
a	0	0	1	0	0
b	0	0	0	0	0
c	0	0	0	0	1
d	1	0	0	0	0
e	1	0	0	1	0

## 2. Adjacency List



# Adjacency Lists vs. Matrix

- **Adjacency Lists**
  - More compact than adjacency matrices if graph has few edges
  - Requires more time to find if an edge exists
- **Adjacency Matrix**
  - Always require  $n^2$  space
    - This can waste a lot of space if the number of edges are sparse
  - Can quickly find if an edge exists

# Directed Graph

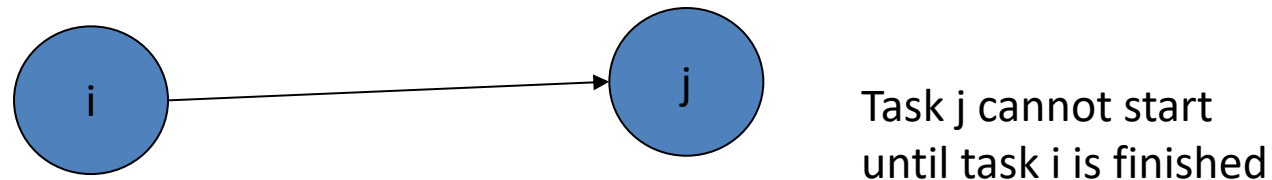
- A graph is directed if direction is assigned to each edge. We call the directed edges *arcs*.
  - An edge is denoted as an ordered pair  $(u, v)$
- Recall: for an undirected graph
  - An edge is denoted  $\{u, v\}$ , which actually corresponds to two arcs  $(u, v)$  and  $(v, u)$

# Directed Acyclic Graph

- A **directed path** is a sequence of vertices  $(v_0, v_1, \dots, v_k)$ 
  - Such that  $(v_i, v_{i+1})$  is an *arc*
- A **directed cycle** is a directed path such that the first and last vertices are the same.
- A directed graph is **acyclic** if it does not contain any directed cycles



- Directed graphs are often **used to represent order-dependent tasks**
- That is we cannot start a task before another task finishes
- We can model this task dependent constraint using *arcs*
- An *arc*  $(i,j)$  means *task j* cannot start until *task i* is finished



- Clearly, for the system not to hang, the graph must be acyclic.

## Breadth-First Search (BFS)

- BFS strategy looks similar to level-order.
- From a given node  $v$ , it first visits itself. Then, it visits every node adjacent to  $v$  before visiting any other nodes.

## Depth-First Search (DFS)

- From a given node  $v$ , it first visits itself. Then, recursively visit its unvisited neighbors one by one.
- Strategy looks similar to pre-order

# BFS Traversal

- ✓ Visit all children of a node, then all grandchildren, etc;
- ✓ Principle is **Queue(LEVEL ORDER)**.

1. Visit  $v$
2. Visit all  $v$ 's neighbors
3. Visit all  $v$ 's neighbors' of neighbors.

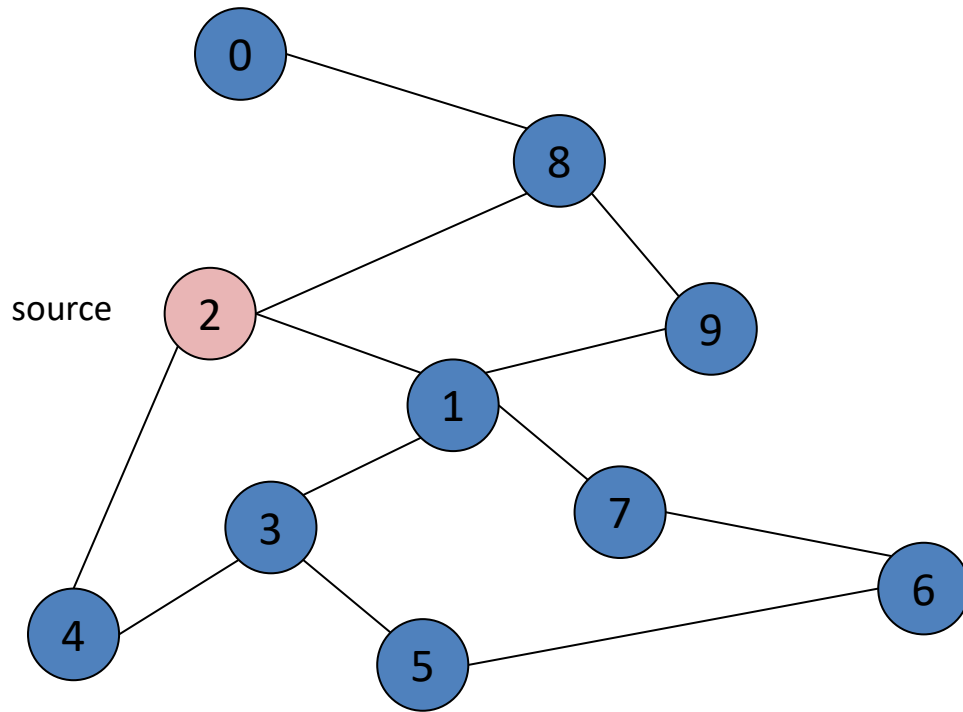
# BFS algorithm

## Algorithm $BFS(s)$

**Input:**  $s$  is the source vertex

**Output:** Mark all vertices that can be visited from  $s$ .

1.   **for** each vertex  $v$
2.       **do**  $flag[v] := \text{false}$ ;
3.    $Q = \text{empty queue}$ ;
4.    $flag[s] := \text{true}$ ;
5.    $enqueue(Q, s)$ ;
6.   **while**  $Q$  is not empty
7.       **do**  $v := dequeue(Q)$ ;
8.       **for** each  $w$  adjacent to  $v$
9.           **do if**  $flag[w] = \text{false}$
10.               **then**  $flag[w] := \text{true}$ ;
11.                $enqueue(Q, w)$



Adjacency List

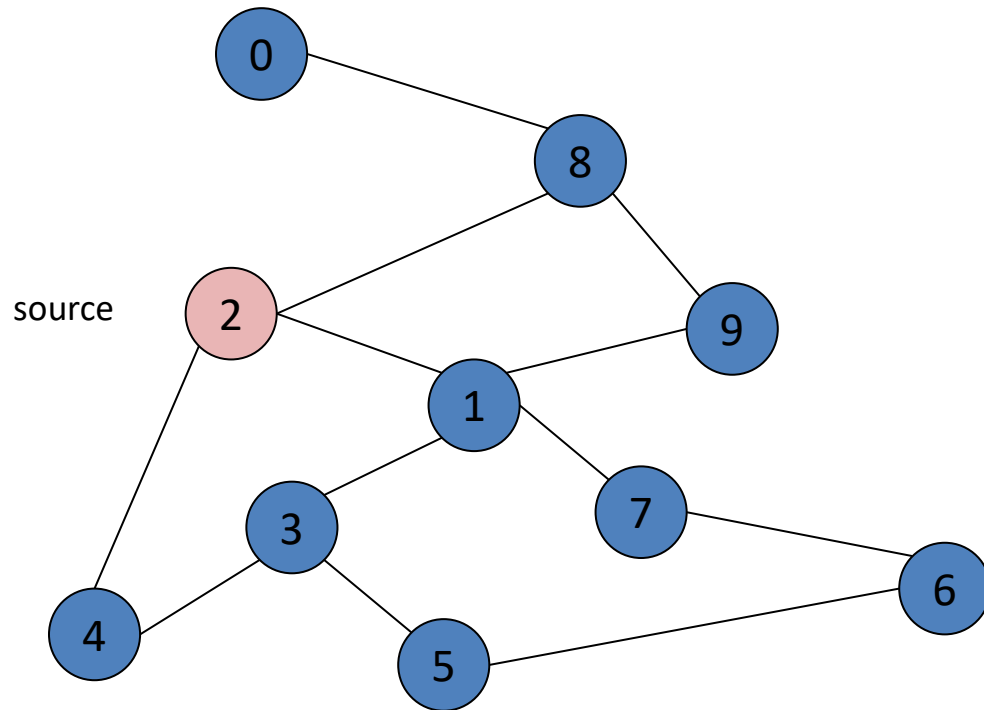
0	8
1	3 7 9 2
2	8 1 4
3	4 5 1
4	2 3
5	3 6
6	7 5
7	1 6
8	2 0 9
9	1 8

Visited Table (T/F)

0	F
1	F
2	F
3	F
4	F
5	F
6	F
7	F
8	F
9	F

Initialize visited  
table (all False)

$Q = \{ \}$  Initialize Q to be empty



## Adjacency List

0	8			
1	3	7	9	2
2	8	1	4	
3	4	5	1	
4	2	3		
5	3	6		
6	7	5		
7	1	6		
8	2	0	9	
9	1	8		

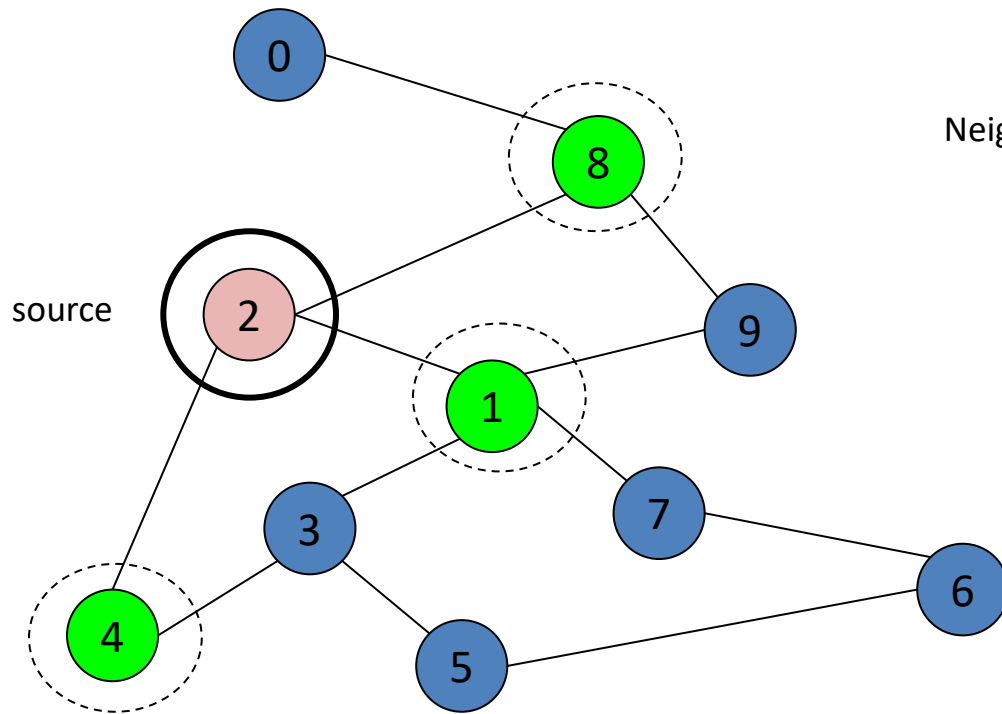
Visited Table (T/F)

0	F
1	F
2	T
3	F
4	F
5	F
6	F
7	F
8	F
9	F

Flag that 2 has been visited.

$$Q = \{ 2 \}$$

Place source 2 on the queue.



$Q = \{2\} \rightarrow \{8, 1, 4\}$

Adjacency List

Neighbors →

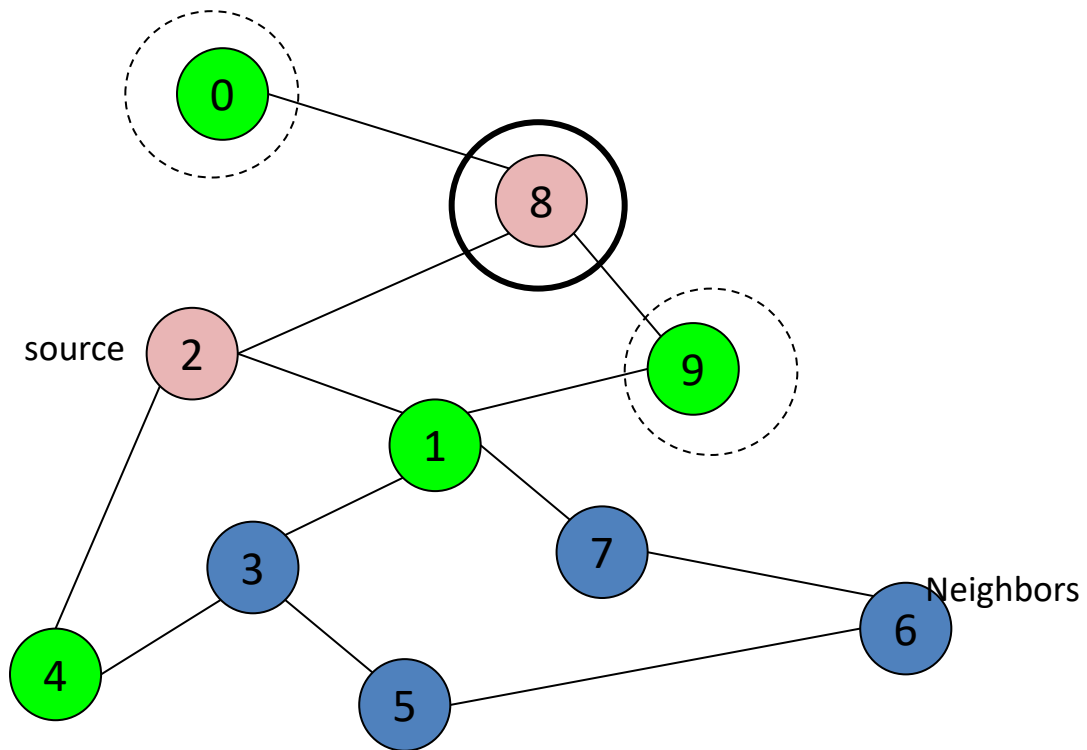
0	8
1	3 7 9 2
2	8 1 4
3	4 5 1
4	2 3
5	3 6
6	7 5
7	1 6
8	2 0 9
9	1 8

Visited Table (T/F)

0	F
1	T
2	T
3	F
4	T
5	F
6	F
7	F
8	T
9	F

Mark neighbors  
as visited.

Dequeue 2. Place all **unvisited** neighbors of 2 on the queue



Adjacency List

0	8
1	3 7 9 2
2	8 1 4
3	4 5 1
4	2 3
5	3 6
6	7 5
7	1 6
8	2 0 9
9	1 8

Visited Table (T/F)

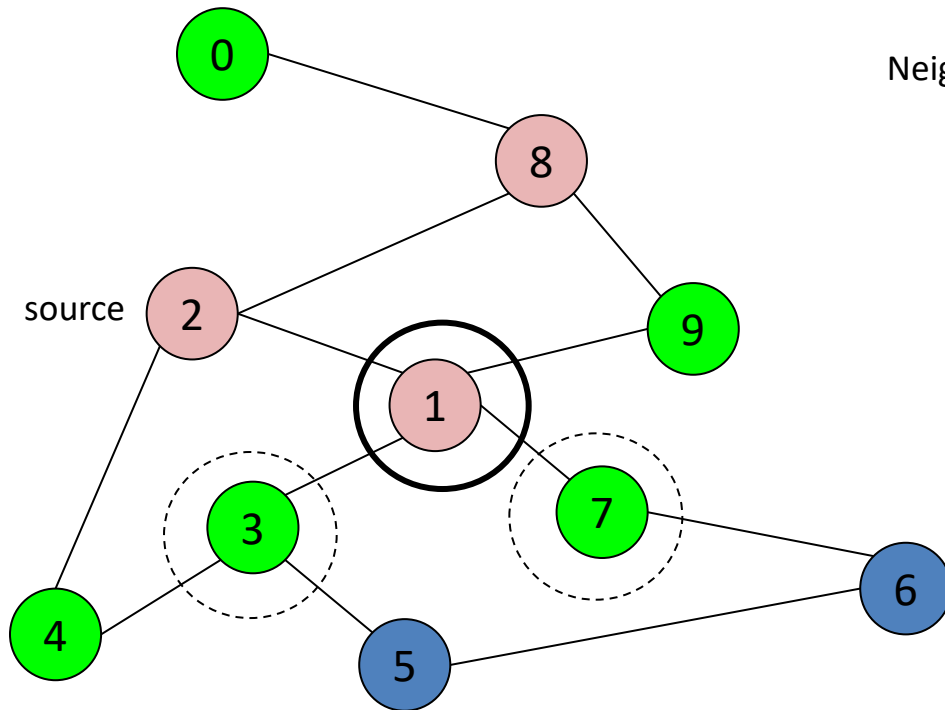
0	T
1	T
2	T
3	F
4	T
5	F
6	F
7	F
8	T
9	T

Mark new visited  
Neighbors.

Q = { 8, 1, 4 } → { 1, 4, 0, 9 }  
Dequeue 8.

- Place all unvisited neighbors of 8 on the queue.
- Notice that 2 is not placed on the queue again, it has been visited!





Adjacency List

Neighbors →

0	8
1	3 7 9 2
2	8 1 4
3	4 5 1
4	2 3
5	3 6
6	7 5
7	1 6
8	2 0 9
9	1 8

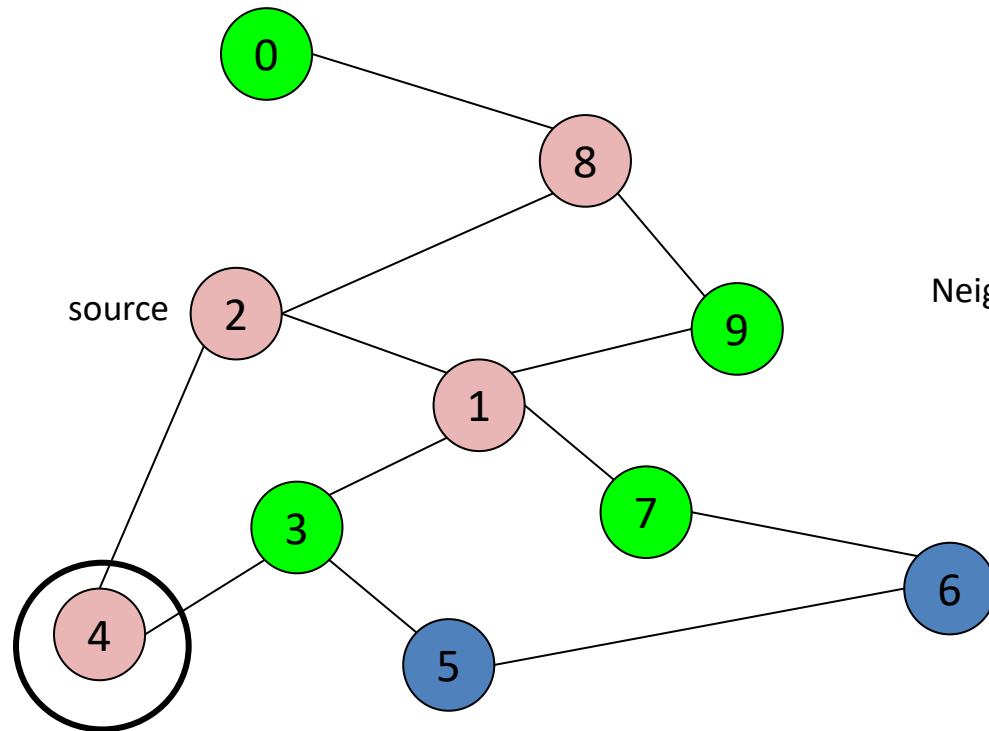
Visited Table (T/F)

0	T
1	T
2	T
3	T
4	T
5	F
6	F
7	T
8	T
9	T

Mark new visited  
Neighbors.

$Q = \{ 1, 4, 0, 9 \} \rightarrow \{ 4, 0, 9, 3, 7 \}$   
Dequeue 1.

- Place all unvisited neighbors of 1 on the queue.
- Only nodes 3 and 7 haven't been visited yet.



Neighbors →

Adjacency List

0	8
1	3 7 9 2
2	8 1 4
3	4 5 1
4	2 3
5	3 6
6	7 5
7	1 6
8	2 0 9
9	1 8

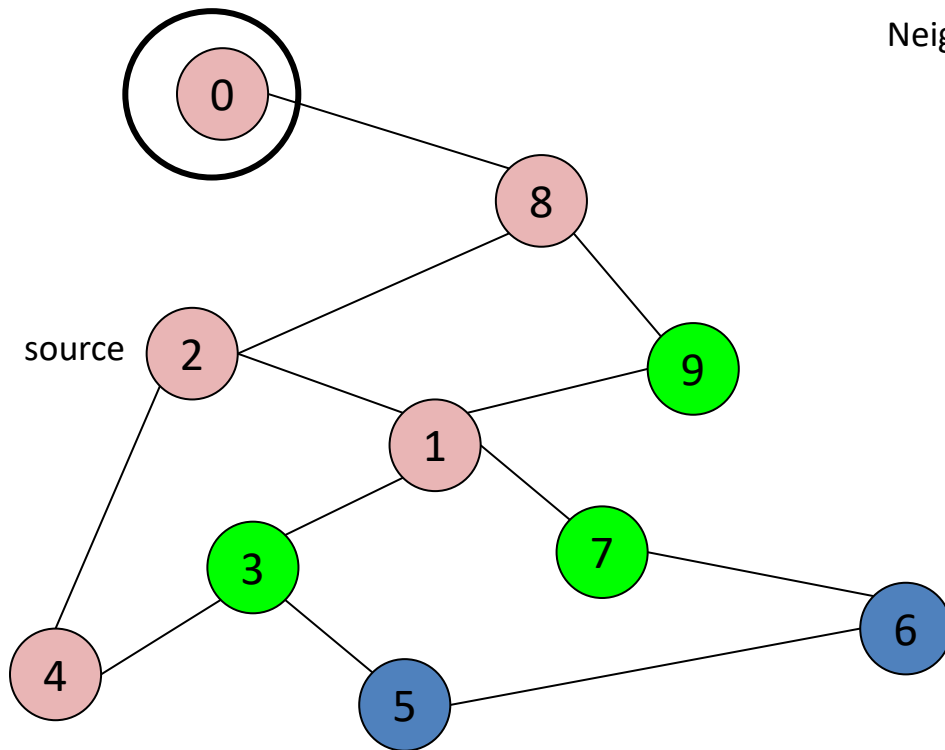
Visited Table (T/F)

0	T
1	T
2	T
3	T
4	T
5	F
6	F
7	T
8	T
9	T

$Q = \{4, 0, 9, 3, 7\} \rightarrow \{0, 9, 3, 7\}$

Dequeue 4.

-- 4 has no unvisited neighbors!



Adjacency List

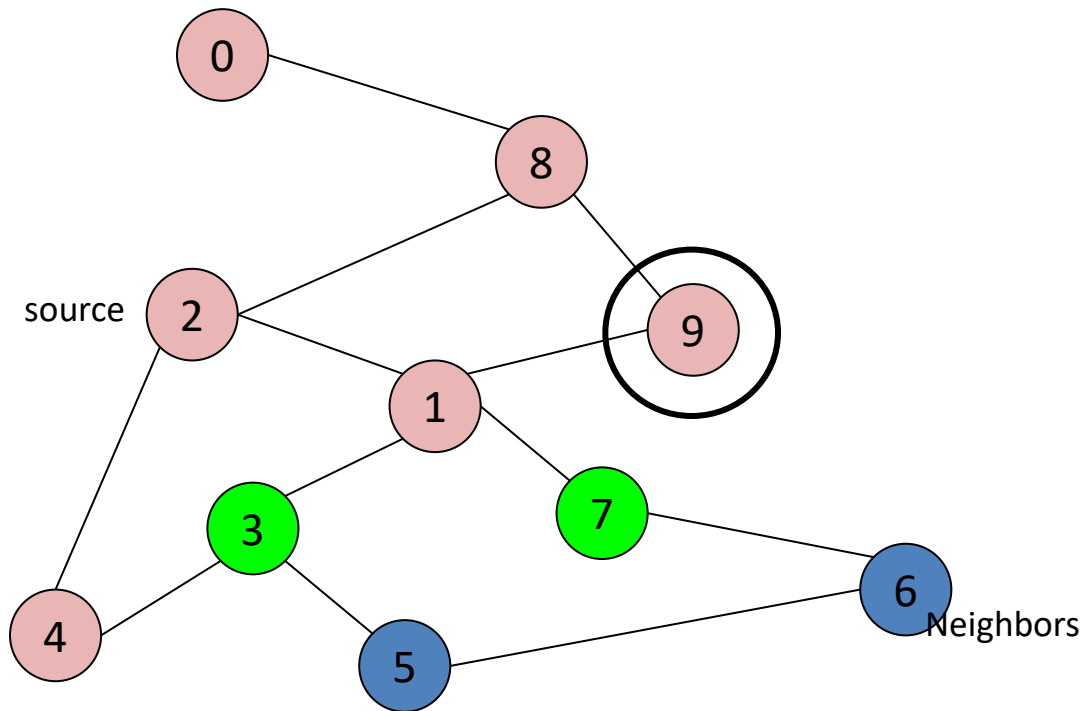
Neighbors →

0	8
1	3 7 9 2
2	8 1 4
3	4 5 1
4	2 3
5	3 6
6	7 5
7	1 6
8	2 0 9
9	1 8

Visited Table (T/F)

0	T
1	T
2	T
3	T
4	T
5	F
6	F
7	T
8	T
9	T

Q = {0, 9, 3, 7} → {9, 3, 7}  
 Dequeue 0.  
 -- 0 has no unvisited neighbors!



Adjacency List

0	8
1	3 7 9 2
2	8 1 4
3	4 5 1
4	2 3
5	3 6
6	7 5
7	1 6
8	2 0 9
9	1 8

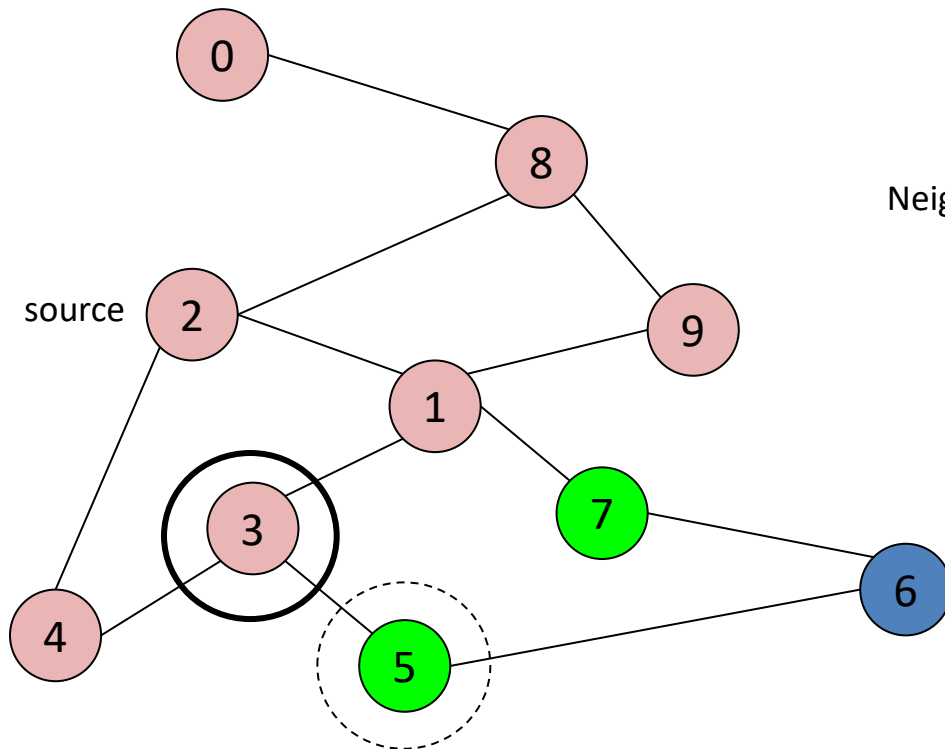
Visited Table (T/F)

0	T
1	T
2	T
3	T
4	T
5	F
6	F
7	T
8	T
9	T

$Q = \{9, 3, 7\} \rightarrow \{3, 7\}$

Dequeue 9.

-- 9 has no unvisited neighbors!



Adjacency List

Neighbors →

0	8
1	3 7 9 2
2	8 1 4
3	4 5 1
4	2 3
5	3 6
6	7 5
7	1 6
8	2 0 9
9	1 8

Visited Table (T/F)

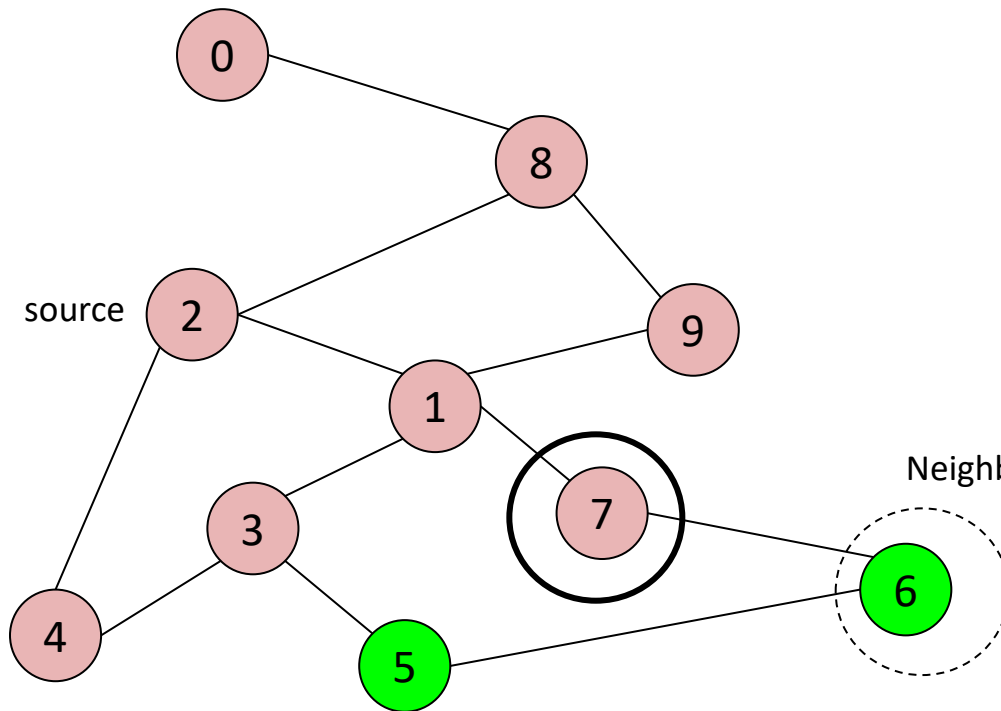
0	T
1	T
2	T
3	T
4	T
5	T
6	F
7	T
8	T
9	T

Mark new visited  
Vertex 5.

$Q = \{3, 7\} \rightarrow \{7, 5\}$

Dequeue 3.

-- place neighbor 5 on the queue.



Adjacency List

0	8
1	3 7 9 2
2	8 1 4
3	4 5 1
4	2 3
5	3 6
6	7 5
7	1 6
8	2 0 9
9	1 8

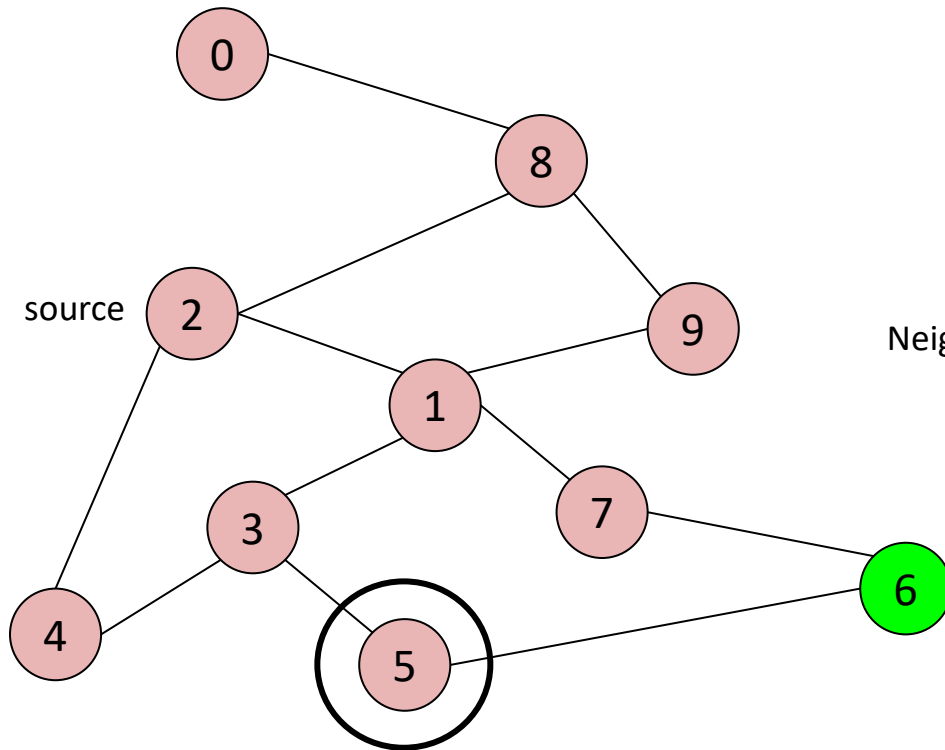
Visited Table (T/F)

0	T
1	T
2	T
3	T
4	T
5	T
6	T
7	T
8	T
9	T

Mark new visited  
Vertex 6.

$Q = \{7, 5\} \rightarrow \{5, 6\}$

Dequeue 7. place neighbor 6 on the queue.



Adjacency List

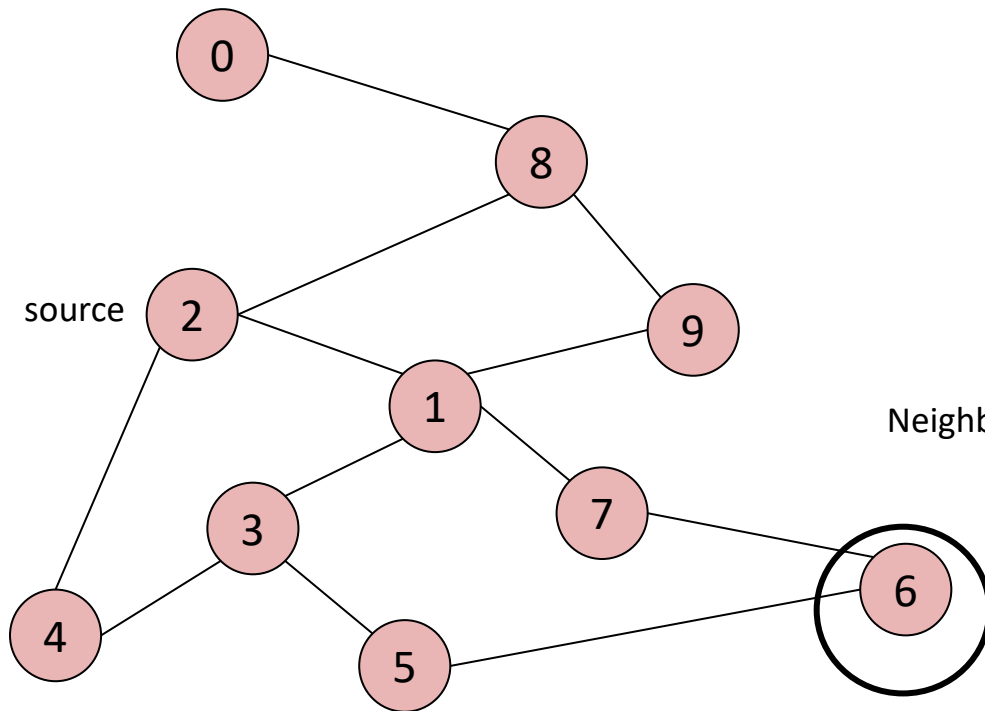
0	8
1	3 7 9 2
2	8 1 4
3	4 5 1
4	2 3
5	3 6
6	7 5
7	1 6
8	2 0 9
9	1 8

Visited Table (T/F)

0	T
1	T
2	T
3	T
4	T
5	T
6	T
7	T
8	T
9	T

$Q = \{5, 6\} \rightarrow \{6\}$

Dequeue 5. No unvisited neighbors of 5.



Adjacency List

0	8
1	3 7 9 2
2	8 1 4
3	4 5 1
4	2 3
5	3 6
6	7 5
7	1 6
8	2 0 9
9	1 8

Visited Table (T/F)

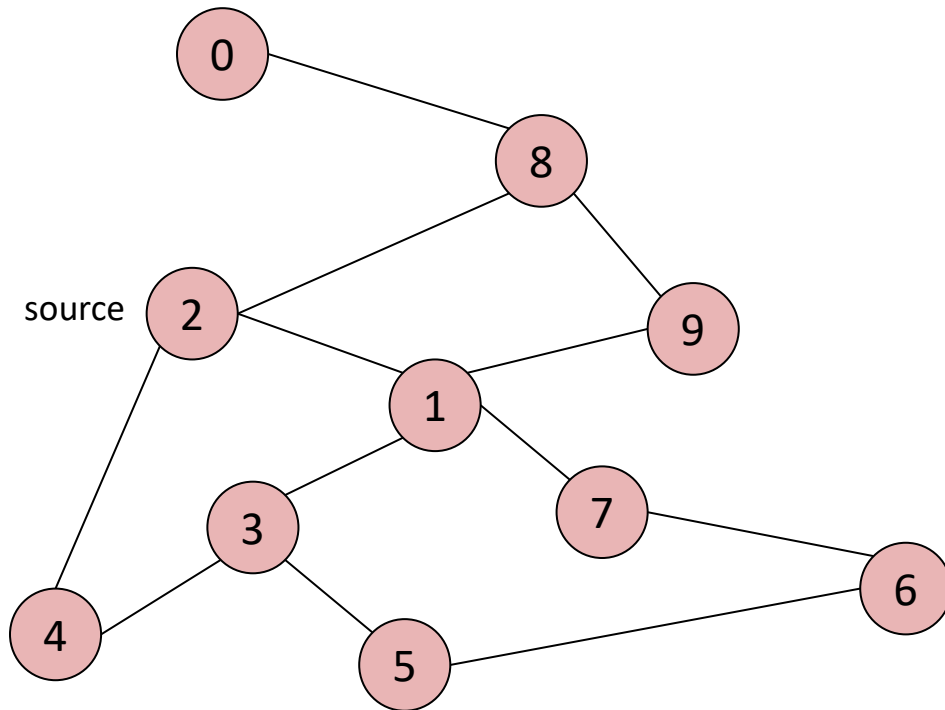
0	T
1	T
2	T
3	T
4	T
5	T
6	T
7	T
8	T
9	T

$Q = \{6\} \rightarrow \{\}$

Dequeue 6. No unvisited neighbors of 6.



# Example



Adjacency List

0	8
1	3 7 9 2
2	8 1 4
3	4 5 1
4	2 3
5	3 6
6	7 5
7	1 6
8	2 0 9
9	1 8

Visited Table (T/F)

0	T
1	T
2	T
3	T
4	T
5	T
6	T
7	T
8	T
9	T

The **BFS traversal** of the given graph is **{2, 8, 1, 4, 0, 9, 3, 7, 5, 6}**

Q = { }      STOP!!! Q is empty!!!

- Time Complexity:

Assume **Adjacency list**

$n$  = number of vertices

$m$  = number of edges

Time Complexity is  **$O(n+m)$**

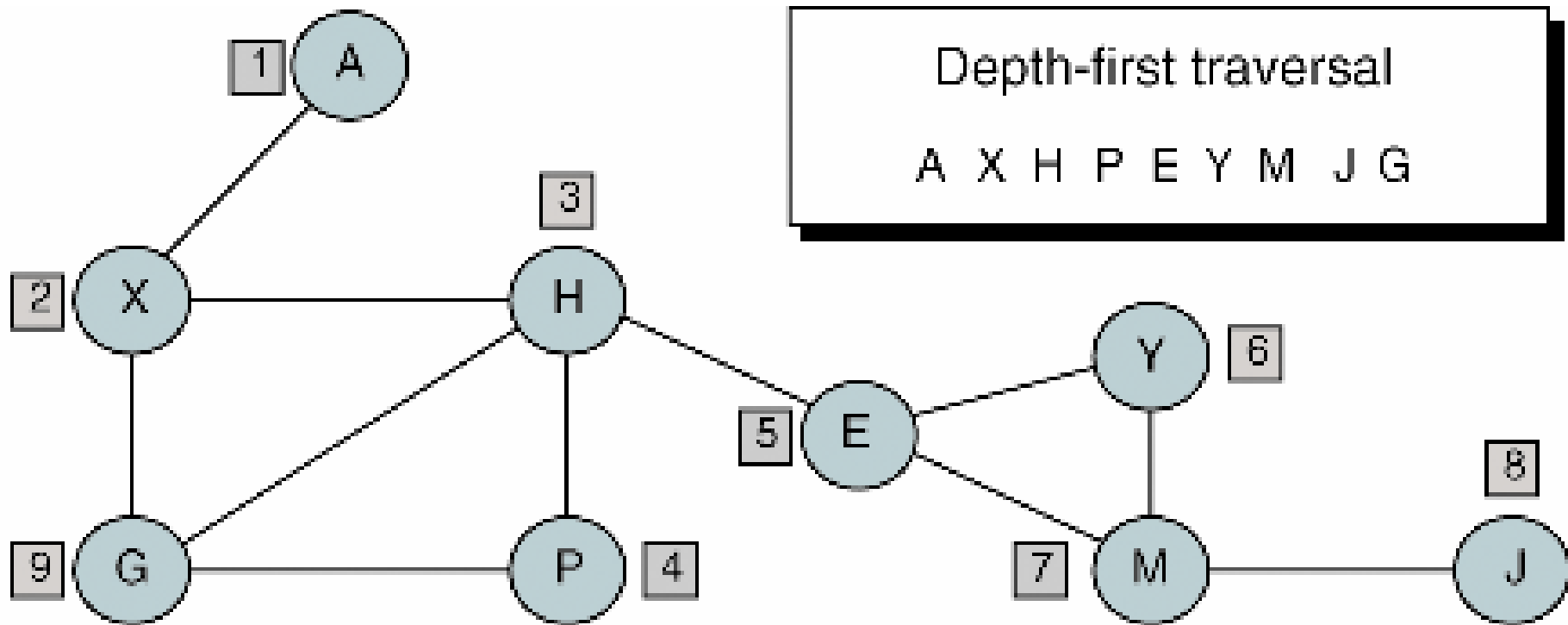
If **Adjacency Matrix** then

Time Complexity is  **$O(n*n)$**

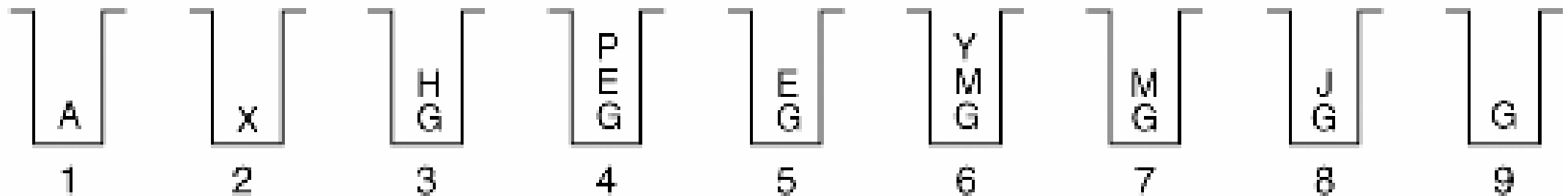
- A recursive algorithm implicitly recording a “backtracking” path from the root to the node currently under consideration
- ✓ Process all of a vertex’s descendants before **we move to an adjacent vertex.**
- ✓ **Visit all descendants of a node, before visiting sibling** nodes
- ✓ Depth first search **use a Stack or LIFO** data structure.
- ✓ Time complexity of is same –  **$O(|V|+|E|)$** . Or  $O(n+m)$

## PROCEDURE:

1. DFS(G,v) ( v is the vertex where the search starts )
2.     Stack S := {}; ( start with an empty stack )
3.     **for** each vertex u, set visited[u] := **false**;
4.     push S, v;
5.     **while** (S is not empty) **do**
6.         u := pop S;
7.         **if** (not visited[u]) then
8.             visited[u] := **true**;
9.             **for** each unvisited neighbour w of u
10.                 push S, w;
11.             end **if**
12.     end **while**
13.    END DFS()



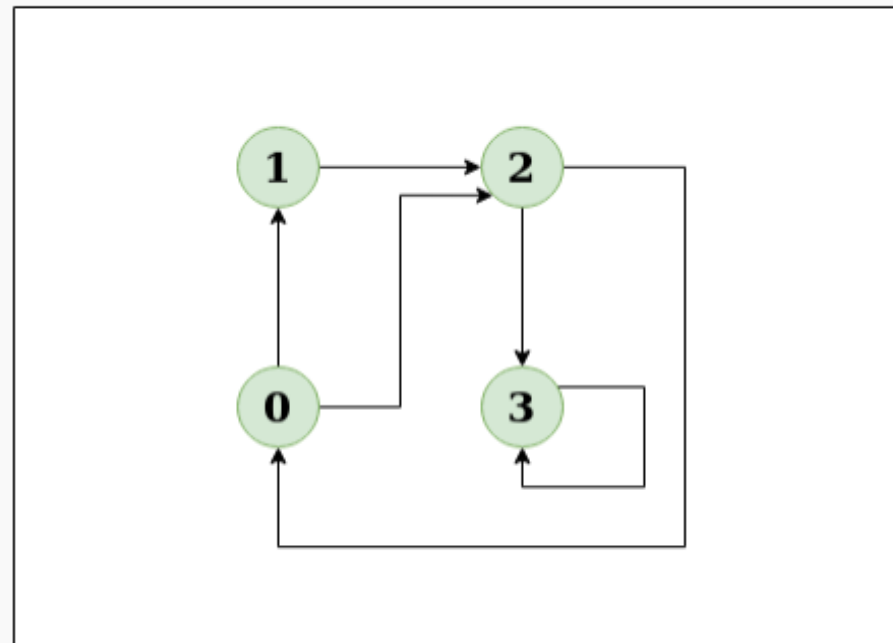
**(a) Graph**



**(b) Stack contents**

The objective is to check whether the graph contains a cycle or not.

*Input:  $N = 4, E = 6$*



*Example of graph*

**Output:** Yes

**Explanation:** The diagram clearly shows a cycle  $0 \rightarrow 2 \rightarrow 0$

*Approach to solve:*

- Use the Depth First Traversal (DFS) technique to find cycle in a directed graph.
- It can be found **only if there is a back edge**.
- **The node points to one of its ancestors**

To detect **Back Edge**

- **Track of the nodes visited** till now and the **nodes** that are in the current recursion stack
- If **during recursion**, we reach a **node** that **is already in the recursion stack**, there is a cycle present in the graph.

## DFS ( current vertex, visited array, recursion stack)

### Algorithm for Directed Graph

#### 1.Initialize Structures:

- Use a **visited array** to track visited nodes.
- Use a **recStack array** (recursive stack) to track nodes currently in the recursion stack.

#### 2.DFS Function:

- Mark the **current node as visited** and **add it to the recStack**.
- For each neighbor of **the current node**:
  - If the neighbor is not visited, recursively call the DFS function.
  - If the neighbor is in the recStack, a **cycle is detected and returns true**.
- Remove the node from the **recStack before returning**.

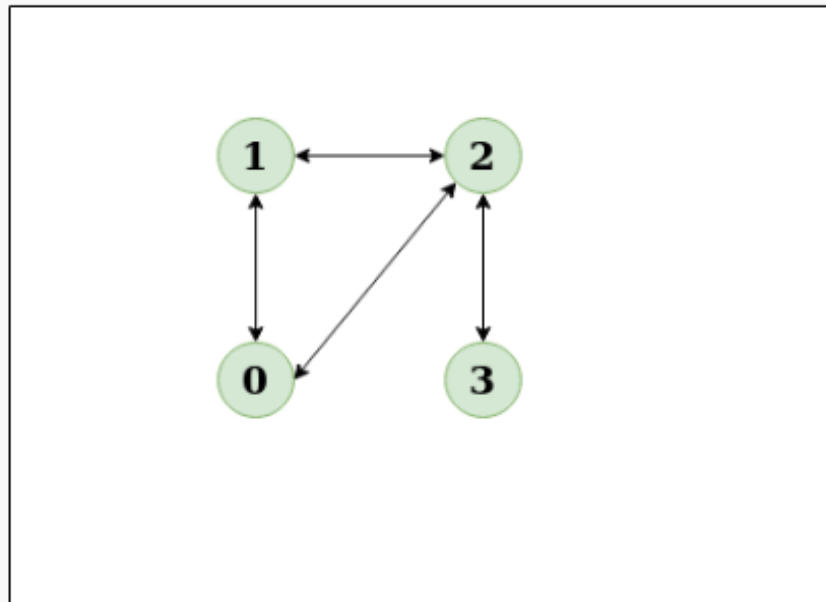
#### 3.Cycle Detection:

- Start DFS from every **unvisited** node in the graph.



The objective is to check whether the graph contains a cycle or not.

*Input:  $N = 4, E = 4$*



*Output: Yes*

*Explanation: The diagram clearly shows a cycle 0 to 2 to 1 to 0*

- Use the Depth First Traversal (DFS) technique to find cycle in a directed graph.
- It can be found **only if there is a back edge**.
- **The node points to one of its ancestors**

To **find the back edge** to any of **its ancestors keep** a visited array and if there is a **back edge to any visited node** then there **is a loop and return true**.

## Algorithm for Undirected Graph

### 1.Initialize Structures:

- Use a **visited array** to track **visited nodes**.

### 2.DFS Function:

- Mark the current node **as visited**.
- For each neighbor of the current node:
  - If the neighbor is not visited, recursively call the DFS function.
  - If the **neighbor is visited** and is **not the parent** of the current node, a **cycle is detected and return true**.

### 3.Cycle Detection:

- Start DFS from every unvisited node in the graph.

- Given a directed graph, **find out if a vertex v is reachable from another vertex u** for all vertex pairs (u, v) in the given graph.
- Reachable means that **there is a path** from **vertex u to v**.
- The reach-ability matrix is called transitive closure of a graph

An  $O(V+E)$  algorithm is proposed to derive the objective

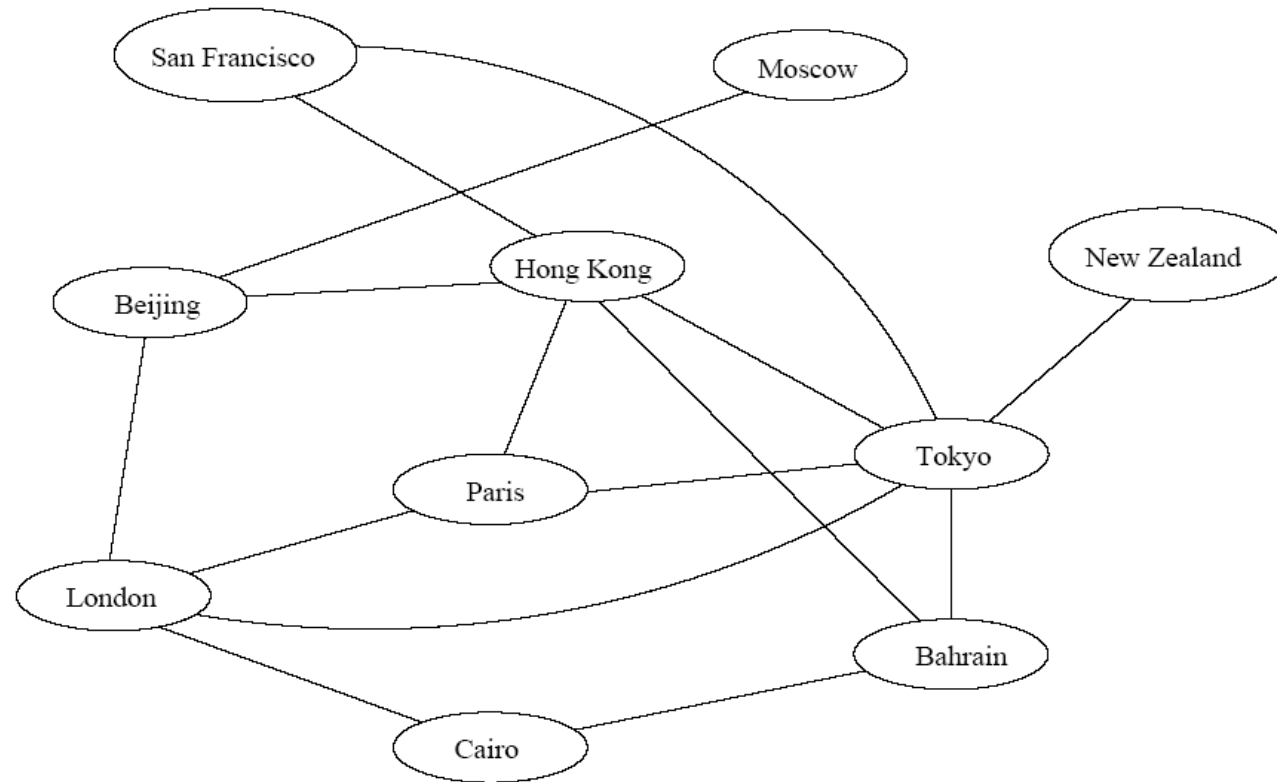
### Steps of the algorithm:

1. Initialize a matrix of size  $N \times N$ , where  $N$  is the number of vertices in the graph. This matrix will **represent the transitive closure** of the graph.
2. For each vertex  $v$  in the graph, perform a DFS starting at  $v$ . During the DFS, **mark all vertices that are reachable from  $v$ .**
3. Once the DFS is complete, **update the matrix to include all the edges** that can be formed by following a **path from  $v$  to each of the reachable vertices.**

Specifically, if **vertex  $i$  is reachable from vertex  $j$**  during the DFS, **mark the entry  $(j,i)$**  in the matrix.

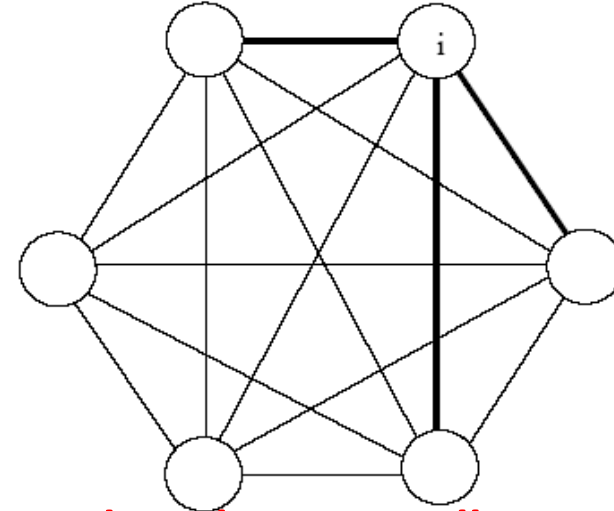
1. Repeat step 2 and 3 for all vertices in the graph.
2. The resulting matrix represents the transitive closure of the graph.

## Air flight system



- Each **vertex represents a city**
- Each **edge represents a direct flight** between two cities
- A query on direct flights becomes a **query on whether an edge exists**
- A **query on how to get to a location** is “**does a path exist from A to B**”
- We can even associate costs to edges (**weighted graphs**), then ask “**what is the cheapest path from A to B**”

## Wireless communication



- A typical **wireless communication problem** is: how to **broadcast between all stations** such that **they are all connected** and **the power consumption is minimized**.
- Can be represented by a **weighted complete graph** (every two vertices are connected by an edge).
- Each **edge represents** the **Euclidean distance**  $d_{ij}$  between two **stations**.
- Each station **uses a certain power**  $i$  to **transmit messages**. Given **this power**  $i$ , only a **few nodes can be reached (bold edges)**. A station reachable by  $i$  then use its own power to relay the message to other stations not reachable by  $i$ .

- ✓ Route Finding
- ✓ Game-Playing
- ✓ Critical Path Analysis
- ✓ Travel arrangement
- ✓ Communication and transportation
- ✓ Networks
- ✓ Logic circuits
- ✓ Computer aided designs
- ✓ Road network
- ✓ Pipeline network
- ✓ Activity chart.
- ✓ Maps, Schedules
- ✓ Computer networks



7.	Trees	Graphs
1.Path	Tree is special form of graph i.e. <b>minimally connected graph</b> and having only one path between any two vertices.	In graph there can be more than one path i.e. graph can have uni-directional or bi-directional paths (edges) between nodes
2.Loops	Tree is a special case of graph having no <b>loops</b> , no <b>circuits</b> and no self-loops.	Graph can have loops, circuits as well as can have <b>self-loops</b> .
3.Root Node	In tree there is exactly one root node and every <b>child</b> have only one <b>parent</b> .	In graph there is no such concept of <b>root</b> node.
4.Parent Child relationship	In trees, there is parent child relationship so flow can be there with direction top to bottom or vice versa.	In Graph there is no such parent child relationship.
5. Complexity	Trees are less complex then graphs as having no cycles, no self-loops and still connected.	Graphs are more complex in compare to trees as it can have cycles, loops etc
6. Types of Traversal	Tree traversal is a kind of special case of traversal of graph. Tree is traversed in <b>Pre-Order, In-Order</b> and <b>Post-Order</b> (all three in DFS or in BFS algorithm)	Graph is traversed by <b>DFS: Depth First Search</b> and in <b>BFS : Breadth First Search algorithm</b>

	<b>Trees</b>	<b>Graphs</b>
<b>7. Connection Rules</b>	In trees, there are many rules / restrictions for making connections between nodes through edges.	In graphs no such rules/ restrictions are there for connecting the nodes through edges.
<b>8. DAG</b>	Trees come in the category of <b>DAG : Directed Acyclic Graphs</b> is a kind of directed graph that have no cycles.	Graph can be <b>Cyclic or Acyclic</b> .
<b>9. Different Types</b>	Different types of trees are : <b>Binary Tree , Binary Search Tree, AVL tree, Heaps.</b>	There are mainly two types of Graphs : <b>Directed and Undirected graphs.</b>
<b>10. Applications</b>	Tree applications : sorting and searching like Tree Traversal & Binary Search.	Graph applications : Coloring of maps, in OR ( <b>PERT &amp; CPM</b> ), algorithms, Graph coloring, job scheduling, etc.
<b>11. No. of edges</b>	Tree always has <b>n-1</b> edges.	In Graph, no. of edges depend on the graph.
<b>12. Model</b>	Tree is a <b>hierarchical model</b> .	Graph is a <b>network model</b> .
<b>13. Figure</b>		

