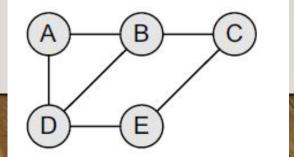
GRAPHS

UNDIRECTED GRAPH

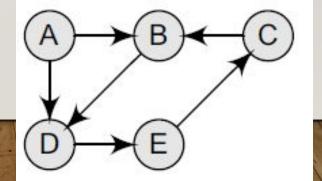
Definition

- A graph G is defined as an ordered set (V, E), where V(G)
 represents the set of vertices and E(G) represents the edges that
 connect these vertices.
- Graph V(G) = {A, B, C, D, E} and E(G) = {(A, B), (B, C), (A, D), (B, D), (D, E), (C, E)}.
- Note that there are five vertices or nodes and six edges in the graph.

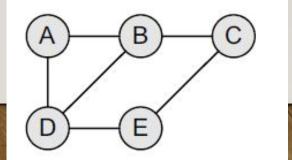


DIRECTED GRAPH

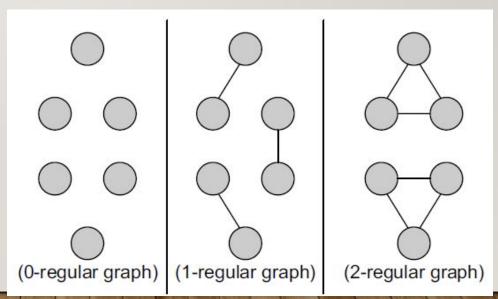
- In an undirected graph, edges do not have any direction associated with them. That is, if an edge is drawn between nodes A and B, then the nodes can be traversed from A to B as well as from B to A.
- In a directed graph, edges form an ordered pair. If there is an edge from A to B, then there is a path from A to B but not from B to A. The edge (A, B) is said to initiate from node A (also known as initial node) and terminate at node B (terminal node).



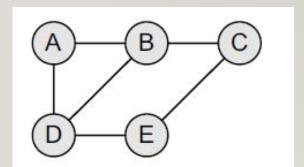
- Adjacent nodes or neighbors For every edge, e = (u, v) that connects nodes u and v, the nodes u and v are the end-points and are said to be the adjacent nodes or neighbours.
- **Degree of a node** Degree of a node u, deg(u), is the total number of edges containing the node u.
- If deg(u) = 0, it means that u does not belong to any edge and such a node is known as an isolated node.



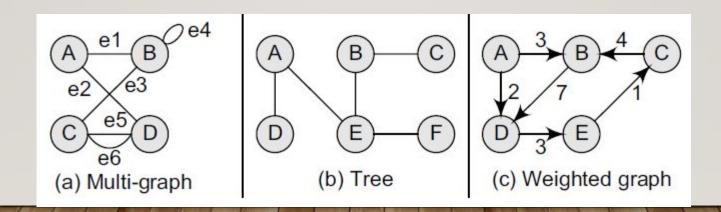
• *Regular graph* It is a graph where each vertex has the same number of neighbors. That is, every node has the same degree. A regular graph with vertices of degree k is called a k-regular graph or a regular graph of degree k.



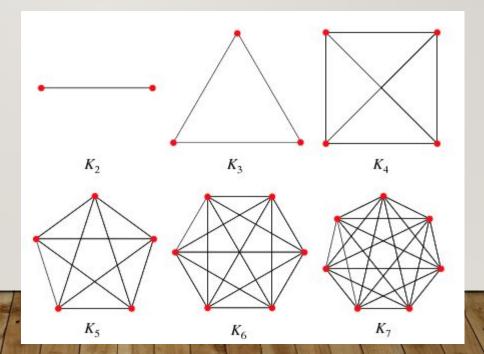
- Path A path P written as P = {v0, v1, v2, ..., vn}, of length n from a node u to v is defined as a sequence of (n+1) nodes. Here, u = v0, v = vn and vi-1 is adjacent to vi for i = 1, 2, 3,..., n.
- *Cycle* A path in which the first and the last vertices are same. A *simple cycle* has no repeated edges or vertices (except the first and last vertices).



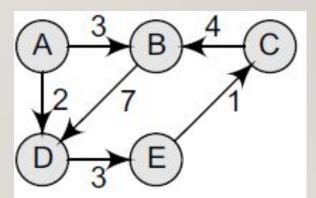
- *Connected graph* A graph is said to be connected if for any two vertices (u, v) in V there is a path from u to v. That is to say that there are no isolated nodes in a connected graph.
- A connected graph that does not have any cycle is called a tree. Therefore, a tree is treated as a special graph.



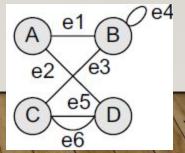
• *Complete graph* A graph G is said to be complete if all its nodes are fully connected. That is, there is a path from one node to every other node in the graph. A complete graph has n(n-1)/2 edges, where n is the number of nodes in G.



• Labelled graph or weighted graph A graph is said to be labelled if every edge in the graph is assigned some data. In a weighted graph, the edges of the graph are assigned some weight or length. The weight of an edge denoted by w(e) is a positive value which indicates the cost of traversing the edge. Figure 13.4(c) shows a weighted graph.



- Multiple edges Distinct edges which connect the same end-points are called multiple edges. That is, e = (u, v) and e' = (u, v) are known as multiple edges of G.
- Loop An edge that has identical end-points is called a loop. That is, e=(u,u).
- *Multi-graph* A graph with multiple edges and/or loops is called a multi-graph.
- Size of a graph The size of a graph is the total number of edges in it.



DIRECTED GRAPHS

- A directed graph G, also known as a digraph, is a graph in which every edge
 has a direction assigned to it. An edge of a directed graph is given as an
 ordered pair (u, v) of nodes in G. For an edge (u, v),
- The edge begins at u and terminates at v.
- u is known as the **origin** or initial point of **e**. Correspondingly, **v** is known as the **destination** or terminal point of **e**.
- u is the predecessor of v. Correspondingly, v is the successor of u.
- Nodes u and v are adjacent to each other.

DIRECTED GRAPH TERMINOLOGY

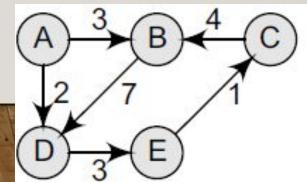
- Out-degree of a node The out-degree of a node u, written as outdeg(u), is the number of edges that originate at u.
- *In-degree of a node* The in-degree of a node u, written as indeg(u), is the number of edges that terminate at u.
- **Degree of a node** The degree of a node, written as deg(u), is equal to the sum of in-degree and out-degree of that node. Therefore,

$$deg(u) = indeg(u) + outdeg(u)$$
.

• *Isolated vertex* A vertex with degree zero. Such a vertex is not an end-point of any edge.

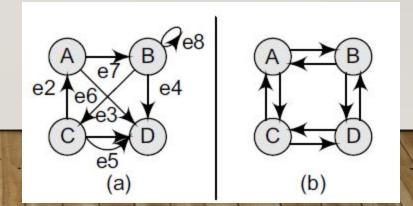
DIRECTED GRAPH TERMINOLOGY

- *Source* A node u is known as a source if it has a positive out-degree but a zero in-degree.
- Sink A node u is known as a sink if it has a positive in-degree but a zero out-degree.
- *Reachability* A node v is said to be reachable from node u, if and only if there exists a (directed) path from node u to node v.
- Strongly connected directed graph A digraph is said to be strongly connected if and only if there exists a path between every pair of nodes in G. That is, if there is a path from node u to v, then there must be a path from node v to u.



DIRECTED GRAPH TERMINOLOGY

- Weakly connected digraph A directed graph is said to be weakly connected if it is connected by ignoring the direction of edges. That is, in such a graph, it is possible to reach any node from any other node by traversing edges in any direction (may not be in the direction they point). The nodes in a weakly connected directed graph must have either out-degree or in-degree of at least 1.
- *Simple directed graph* A directed graph G is said to be a simple directed graph if and only if it has no parallel edges. However, a simple directed graph may contain cycles



APPLICATIONS

- Graphs are widely used to model any situation where entities or things are related to each other in pairs. For example, the following information can be represented by graphs:
 - Family trees in which the member nodes have an edge from parent to each of their children.
 - *Transportation networks* in which nodes are airports, intersections, ports, etc. The edges can be airline flights, one-way roads, shipping routes, etc.

REPRESENTATION OF GRAPHS

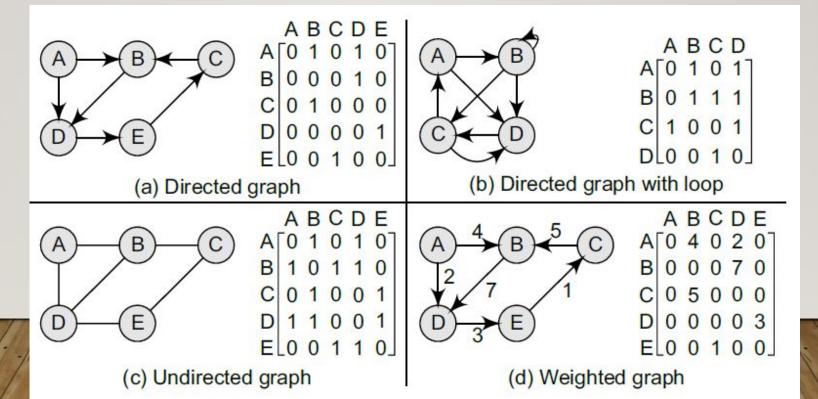
REPRESENTATION OF GRAPHS

Three common ways of storing graphs:

- 1. Sequential representation by using an adjacency matrix.
- 2. Linked representation by using an adjacency list that stores the neighbors of a node using a linked list.
- 3. Adjacency multi-list which is an extension of linked representation.

ADJACENCY MATRIX

In an adjacency matrix, the rows and columns are labelled by graph vertices. An entry aij in the adjacency matrix will contain 1, if vertices vi and vj are adjacent to each other. However, if the nodes are not adjacent, aij will be set to zero.

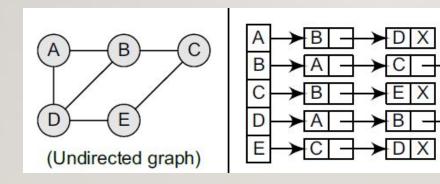


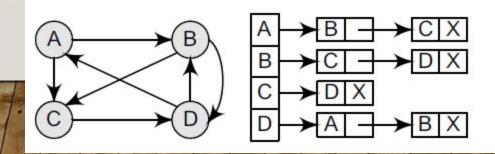
OBSERVATIONS

- The adjacency matrix of an undirected graph is symmetric.
- The memory use of an adjacency matrix is $O(n^2)$, where n is the number of nodes in the graph.
- Number of 1s (or non-zero entries) in an adjacency matrix is equal to the number of edges in the graph.
- The adjacency matrix for a weighted graph contains the weights of the edges connecting the nodes.

ADJACENCY LIST REPRESENTATION

• This structure consists of a list of all nodes in G. Furthermore, every node is in turn linked to its own list that contains the names of all other nodes that are adjacent to it.

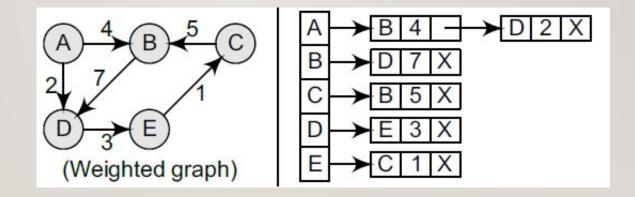




► E X

ADJACENCY LIST REPRESENTATION FOR WEIGHTED GRAPH

• Adjacency lists can also be modified to store weighted graphs



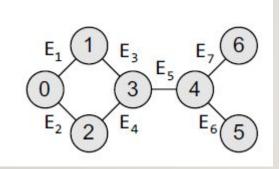
OBSERVATIONS

- The adjacency matrix of an undirected graph is symmetric.
- It is easy to follow and clearly shows the adjacent nodes of a particular node.
- It is often used for storing graphs that have a small-to-moderate number of edges. That is, an adjacency list is preferred for representing sparse graphs in the computer's memory; otherwise, an adjacency matrix is a good choice.
- Adding new nodes in G is easy and straightforward when G is represented using an adjacency list.
- Adding new nodes in an adjacency matrix is a difficult task, as the size of the matrix needs to be changed and existing nodes may have to be reordered.

ADJACENCY MULTI-LIST REPRESENTATION

- Adjacency multi-list is an edge-based rather than a vertex-based representation of graphs. A multi-list representation basically consists of two parts:
 - a directory of nodes' information
 - a set of linked lists storing information about edges.

ADJACENCY MULTI-LIST REPRESENTATION



Edge 1	0	1	Edge 2	Edge 3
Edge 2	0	2	NULL	Edge 4
Edge 3	1	3	NULL	Edge 4
Edge 4	2	3	NULL	Edge 5
Edge 5	3	4	NULL	Edge 6
Edge 6	4	5	Edge 7	NULL
Edge 7	4	6	NULL	NULL

GRAPH TRAVERSAL ALGORITHMS

GRAPH TRAVERSAL ALGORITHMS

- There are two standard methods of graph traversal. These two methods are:
 - I. Breadth-first search
 - uses QUEUE
 - 2. Depth-first search
 - uses STACK

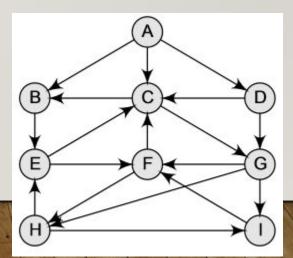
STATUS OF VERTEX IN ALGORITHMS

Value of status and its significance

Status	State of the node	Description
1	Ready	The initial state of the node N
2	Waiting	Node N is placed on the queue or stack and waiting to be processed
3	Processed	Node N has been completely processed

BREADTH-FIRST SEARCH ALGORITHM (BFS)

- Breadth-first search (BFS) is a graph search algorithm that begins at the root node and explores all the neighboring nodes.
- Then for each of those nearest nodes, the algorithm explores their unexplored neighbor nodes, and so on, until it finds the goal

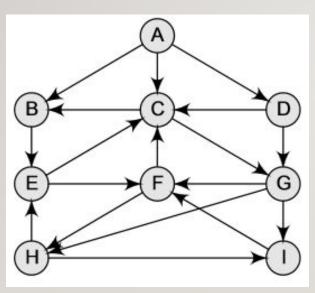


BREADTH-FIRST SEARCH ALGORITHM (BFS)

- Step I: SET STATUS=I (ready state) for each node in G
- Step 2: Enqueue the starting node A and set its STATUS=2(waiting state)
- Step 3: Repeat Steps 4 and 5 until QUEUE is empty
- Step 4: Dequeue a node N. Process it and set its STATUS=3(processed state).
- Step 5: Enqueue all the neighbors of N that are in the ready state(whose STATUS=I) and set their STATUS=2(waiting state)[END OF LOOP]
- Step 6: EXIT

BREADTH-FIRST SEARCH ALGORITHM (BFS)

Initialization



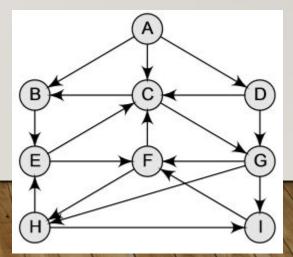
	A	В	C	D	E	F	G	Н	I
		1	1	1					
)					1				
		1					1		
			1				1		
			1			1			
			1					1	
						1		1	1
7					1				1
						1			

Α	7.7				200			_
1	1	1	1	1	1	1	1	1

QL	JEU	E	-20	250	1001	100	200 20
					341 145	1947	î

DEPTH-FIRST SEARCH ALGORITHM (DFS)

- The depth-first search algorithm progresses by expanding the starting node of G and then going deeper and deeper until the goal node is found, or until a node that has no children is encountered.
- When a dead-end is reached, the algorithm backtracks, returning to the most recent node that has not been completely explored



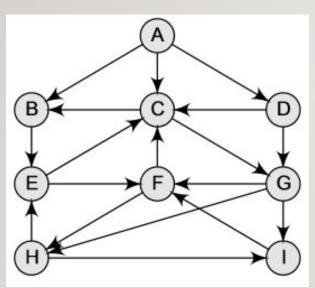
DEPTH-FIRST SEARCH ALGORITHM (DFS)

- Step I: SET STATUS=I (ready state) for each node in G
- Step2: Push the starting node A on the stack and set its STATUS=2(waiting state)
- Step3: Repeat Steps 4 and 5 until STACK is empty
- Step4: Pop the top node N. Process it and set its STATUS=3(processed state)
- Step5: Push on the stack all the neighbors of N that are in the ready state (whose STATUS=1) and set their STATUS=2(waiting state)[END OF LOOP]Step
- 6: EXIT

DEPTH-FIRST SEARCH ALGORITHM (DFS)

STACK

Initialization



Α		_						_
1	1	1	1	1	1	1	1	1

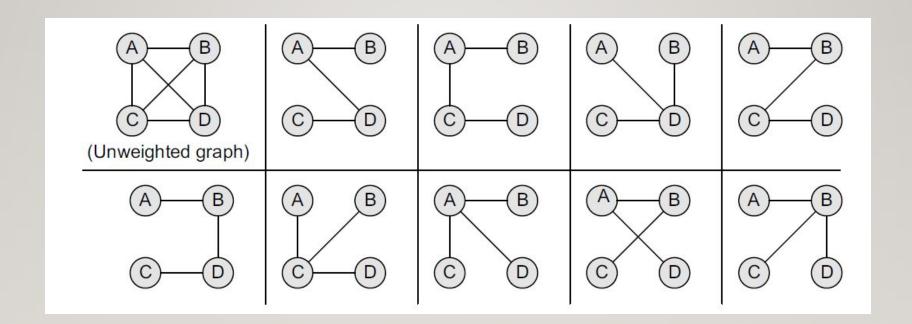
	Α	В	C	D	E	F	G	Н	ı
Α		1	1	1					
B C					1				
С		1					1		
D			1				1		
E			1			1			
F G			1					1	
G						1		1	1
Н					1				1
l						1			

SHORTEST PATH ALGORITHM

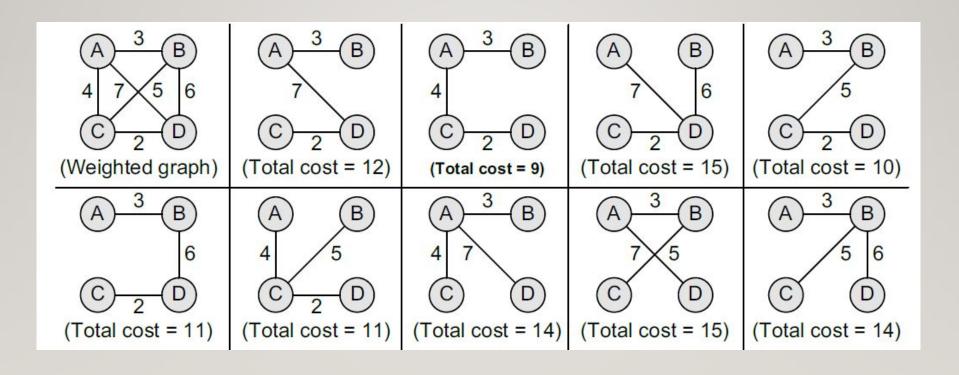
MINIMUM SPANNING TREE

- A spanning tree of a connected, undirected graph
 G is a sub-graph of G which is a tree that
 connects all the vertices together
- A graph G can have many different spanning trees
- A minimum spanning tree (MST) is defined as a spanning tree with weight less than or equal to the weight of every other spanning tree

MINIMUM SPANNING TREE



MINIMUM SPANNING TREE



PROPERTIES OF MST

- Possible multiplicity: There can be multiple minimum spanning trees of the same weight.
- **Uniqueness:** When each edge in the graph is assigned a different weight, then there will be only one unique minimum spanning tree.
- Minimum-cost subgraph: If the edges of a graph are assigned non-negative weights, then a
 minimum spanning tree is in fact the minimum-cost subgraph or a tree that connects all
 vertices.
- Cycle property: If there exists a cycle C in the graph G that has a weight larger than that of other edges of C, then this edge cannot belong to an MST.
- **Usefulness:** Minimum spanning trees can be computed quickly and easily to provide optimal solutions. These trees create a sparse subgraph that reflects a lot about the original graph.
- **Simplicity** The minimum spanning tree of a weighted graph is nothing but a spanning tree of the graph which comprises of n-I edges of minimum total weight.

APPLICATIONS OF MINIMUM SPANNING TREES

- I. MSTs are widely used for designing networks. For instance, people separated by varying distances wish to be connected together through a telephone network. A minimum spanning tree is used to determine the least costly paths with no cycles in this network, thereby providing a connection that has the minimum cost involved.
- 2. **MSTs** are used to find shortest path. While the vertices in the graph denote cities, edges represent the routes between places. More the distance between the cities, higher will be the amount charged. Therefore, MSTs are used to optimize airline routes by finding the least costly path with no cycles.

 Prim's algorithm is a greedy algorithm that is used to form a minimum spanning tree for a connected weighted undirected graph

- *Tree vertices* Vertices that are a part of the minimum spanning tree T.
- *Fringe vertices* Vertices that are currently not a part of T, but are adjacent to some tree vertex.
- Unseen vertices Vertices that are neither tree vertices nor fringe vertices fall under this category.

Step 1: Select a starting vertex

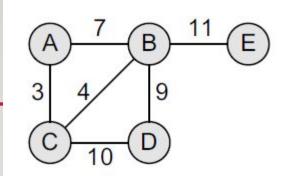
Step 2: Repeat Steps 3 and 4 until there are fringe vertices

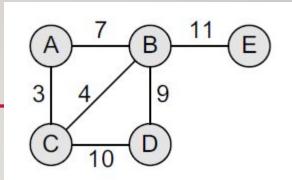
Step 3: Select an edge e connecting the tree vertex and

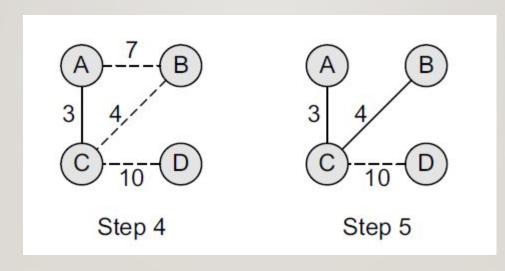
fringe vertex that has minimum weight

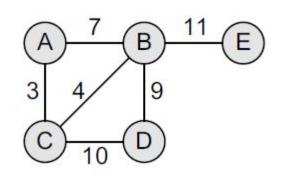
Step 4: Add the selected edge and the vertex to the minimum spanning tree T

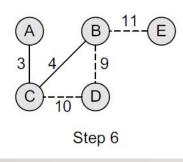
Step 5: EXIT

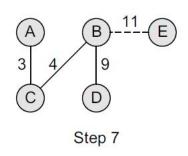


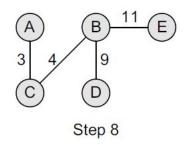






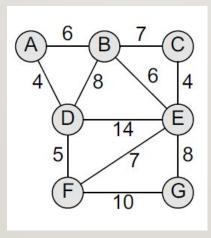


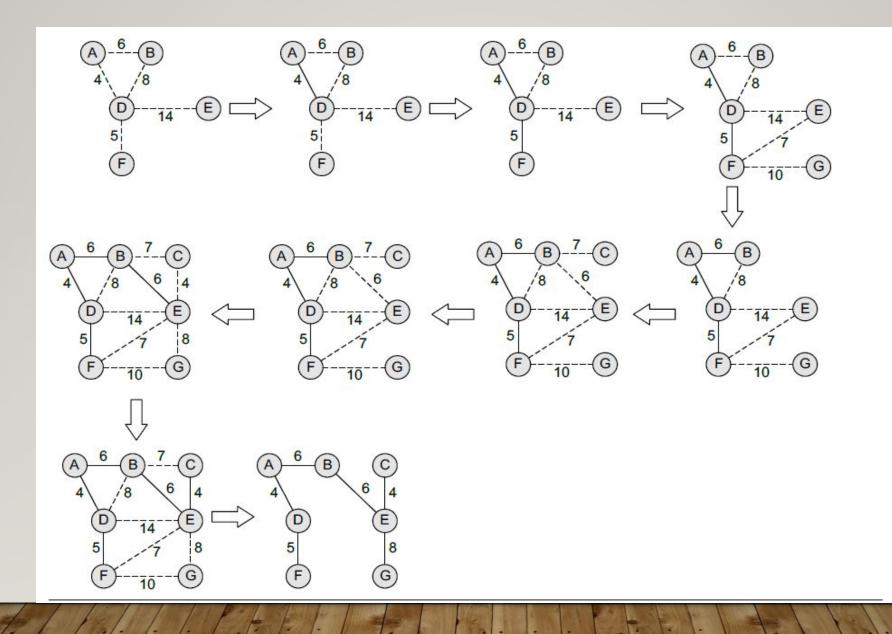




DO IT YOURSELF

 Construct a minimum spanning tree of the graph given in Figure Start the Prim's algorithm from vertex D





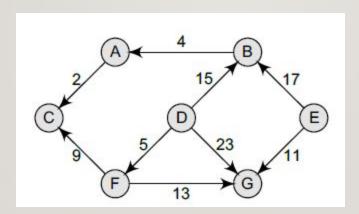
DIJKSTRA'S ALGORITHM

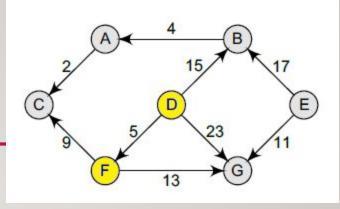
- Dijkstra's algorithm, given by a Dutch scientist Edsger Dijkstra in 1959, is used to find the shortest path tree.
- This algorithm is widely used in network routing protocols

DIJKSTRA'S ALGORITHM

- I. Select the source node also called the initial node
- 2. Define an empty set N that will be used to hold nodes to which a shortest path has been found.
- 3. Label the initial node with 0, and insert it into N.
- 4. Repeat Steps 5 to 7 until the destination node is in N or there are no more labelled nodes in N.
- 5. Consider each node that is not in N and is connected by an edge from the newly inserted node.
- 6. (a) If the node that is not in N has no label then SET the label of the node = the label of the newly inserted node + the length of the edge.
- (b) Else if the node that is not in N was already labelled, then SET its new label = minimum (label of newly inserted vertex + length of edge, old label)
- 7. Pick a node not in N that has the smallest label assigned to it and add it to N.

• Consider the graph G given in figure below, D is the initial node, execute the Dijkstra's algorithm on it.





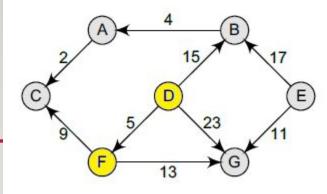
• Step 1: Set the label of D = 0 and $N = \{D\}$.

Vertex	A	В	С	D	E	F	G
Weight				0			
Previous Vertex				-			

• Step 2: Label of D = 0, B = 15, G = 23, and F = 5.

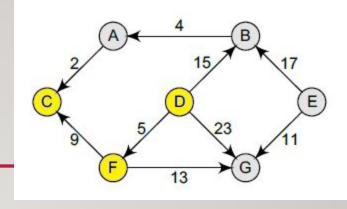
Vertex	A	В	C	D	E	F	G
Weight		15		0		5	23
Previous Vertex		D		-		D	D

Therefore, $N = \{D, F\}$.



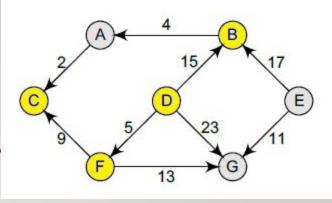
Label of D = 0, B = 15, G has been re-labelled 18 because minimum (5 + 13, 23) = 18, C has been re-labelled 14 (5 + 9). Therefore, N = {D, F, C}.

Vertex	A	В	С	Ι)	E	F	G
Weight		15	14	()		5	18
Previous Vertex		D	F	-	-		D	F



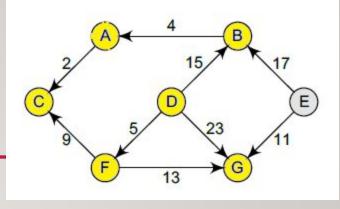
Step 4: Label of D = 0, B = 15, G = 18. Therefore, N = {D, F, C, B}.

Vertex	A	В	С	D	E	F	G
Weight		15	14	0		5	18
Previous Vertex		D	F	-		D	F



Step 5: Label of D = 0, B = 15, G = 18 and A = 19 (15 + 4).
 Therefore, N = {D, F, C, B, G}.

Vertex	A	В	C	D	E	F	G
Weight	19	15	14	0		5	18
Previous Vertex	В	D	F	-		D	F



Step 6: Label of D = 0 and A = 19. Therefore, N = {D, F, C, B, G, A}

A	В	C	D	E	F	G
19	15	14	0		5	18
В	D	F	-		D	F

Note that we have no labels for node E; this means that E
is not reachable from D. Only the nodes that are in N are
reachable from B.

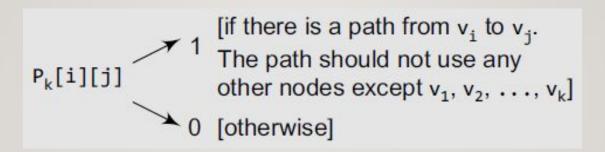
OBSERVATIONS

Prim's	Dijkstra's
Minimum spanning tree algorithm is used to traverse a graph in the most efficient manner	calculates the distance from a given vertex to every other vertex in the graph
Prim's algorithm stores a minimum cost edge	Dijkstra's algorithm stores the total cost from a source node to the current node
Prim's algorithm stores at most one minimum cost edge	Dijkstra's algorithm is used to store the summation of minimum cost edges

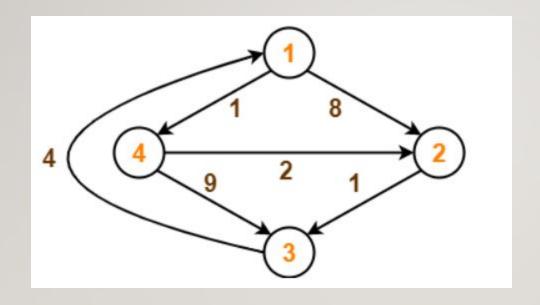
Both the algorithms begin at a specific node and extend outward within the graph, until all other nodes in the graph have been reached

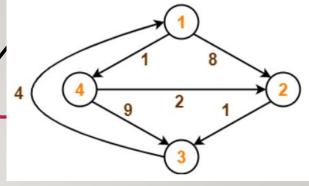
- Used to find Transitive Closure of Matrix
- Finds the length of shortest path between all pair of vertices

• Warshall's algorithm defines matrices P0, P1, P2, . . . , Pn

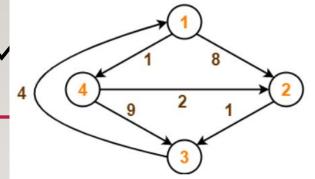


- if P0[i][j] = I, then there exists an edge from node vi to vj.
- If PI[i][j] = I, then there exists an edge from vi to vj that does not use any other vertex except vI.
- If P2[i][j] = I, then there exists an edge from vi to vj that does not use any other vertex except vI and v2.



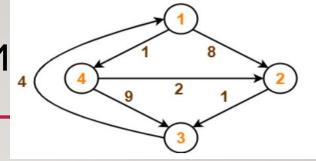


- Write the initial distance matrix.
- It represents the distance between every pair of vertices in the form of given weights.
- For diagonal elements (representing self-loops), distance value =
 0.
- For vertices having a direct edge between them, distance value
 weight of that edge.
- For vertices having no direct edge between them, distance value $= \infty$.



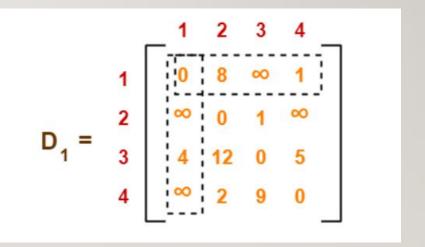
Initial distance matrix for the given graph is-

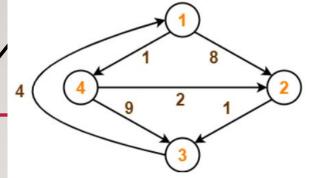
$$D_0 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 8 & \infty & 1 \\ \infty & 0 & 1 & \infty \\ 4 & \infty & 0 & \infty \\ 4 & \infty & 2 & 9 & 0 \end{bmatrix}$$



$$D[i,j] = min(D[i,j],D[i,k]+D[k,j])$$

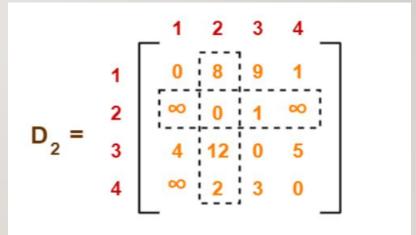
$$D_0 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 8 & \infty & 1 \\ \infty & 0 & 1 & \infty \\ 3 & 4 & \infty & 0 & \infty \\ 4 & \infty & 2 & 9 & 0 \end{bmatrix}$$

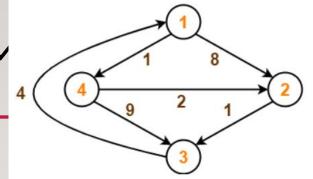




$$D[i,j] = min(D[i,j],D[i,k]+D[k,j])$$

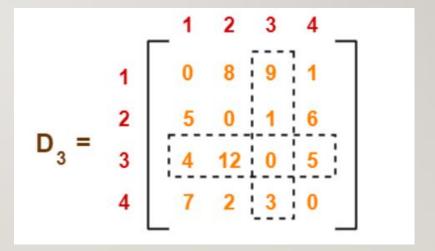
$$D_{1} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 8 & \infty & 1 \\ 2 & 0 & 1 & \infty \\ 4 & 12 & 0 & 5 \\ \infty & 2 & 9 & 0 \end{bmatrix}$$

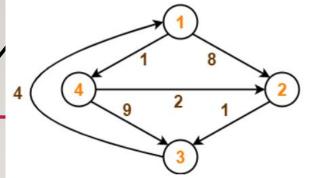




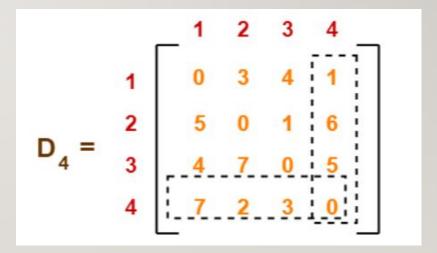
$$D[i,j] = min(D[i,j],D[i,k]+D[k,j])$$

$$D_{2} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 8 & 9 & 1 \\ \hline 0 & 0 & 1 & \infty \\ 4 & 12 & 0 & 5 \\ \hline 0 & 2 & 3 & 0 \end{bmatrix}$$





$$D[i,j] = min(D[i,j],D[i,k]+D[k,j])$$



ALGORITHM

```
Create a |V| x |V| matrix // It represents the distance between every pair of vertices as given
For each cell (i,j) in M do-
   if i = = j
     M[i][j] = 0 // For all diagonal elements, value = 0
   if (i, j) is an edge in E
     M[i][j] = weight(i,j) // If there exists a direct edge between the vertices, value = weight of edge
   else
        M[i][j] = infinity // If there is no direct edge between the vertices, value = \infty
for k from I to |V|
     for i from I to |V|
          for j from I to |V|
                if M[i][j] > M[i][k] + M[k][j]
                      M[i][j] = M[i][k] + M[k][j]
```

Time Complexity-

Floyd Warshall Algorithm consists of three loops over all the nodes. The inner most loop consists of only constant complexity operations. Hence, the asymptotic complexity of Floyd Warshall algorithm is $O(n^3)$. Here, n is the number of nodes in the given graph.

Floyd Warshall Algorithm is best suited for <u>dense</u>
 graphs. This is because its complexity depends only on the
 number of vertices in the given graph.

