Data Structures Instructor: Haifz Tayyeb Javed Week-14-Lecture-01

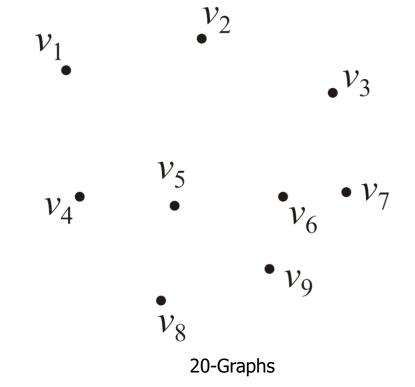
20. Graphs

Graphs

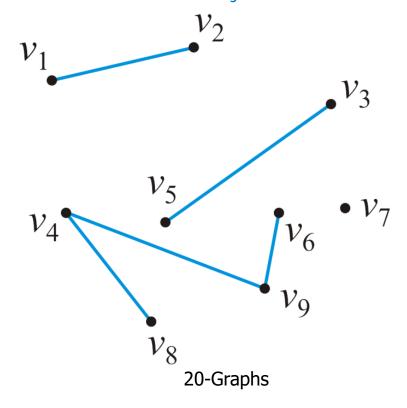
Consider this collection of vertices

$$- V = \{V_1, V_2, \dots, V_9\}$$

- Where |V| = n

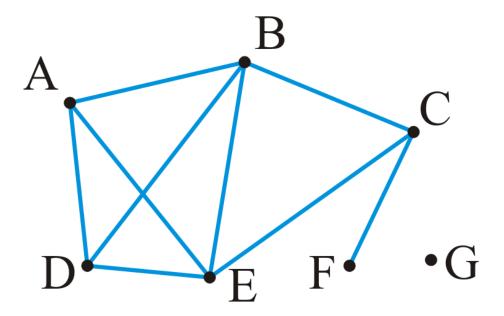


- Associated with these vertices are | E | = 5 edges
 - $E = \{\{v_1, v_2\}, \{v_3, v_5\}, \{v_4, v_8\}, \{v_4, v_9\}, \{v_6, v_9\}\}$
- Pair $\{v_i, v_k\}$ indicates following relations
 - Vertex v_i is adjacent to vertex v_k
 - Vertex v_k is adjacent to vertex v_i



Graphs – Example

• Given |V| = 7 vertices and |E| = 9 edges



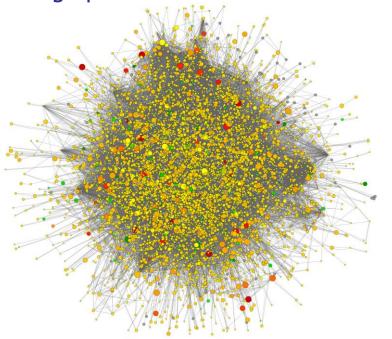
Applications Of Graphs

- Driving Map
 - Vertex = Intersection, destinations
 - Edge = Road
- Airline Traffic
 - Vertex = Cities serviced by the airline
 - Edge = Flight exists between two cities
- Computer networks
 - Vertex = Server nodes, end devices, routers
 - Edge = Data link

Applications Of Graphs

- Many real-world applications concern large graphs
- Web document graph 1 trillion webpages
 - Vertex = Webpage
 - Edge = Hyperlink
- Social networks 2.23 billion users
 - Vertex = Users
 - Edge = Friendship relation

A graph of web links on the internet



Undirected Graphs – Definition

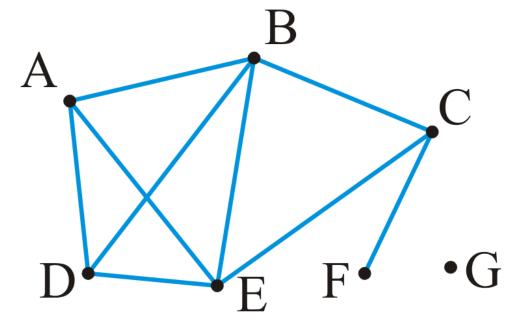
- An undirected Graph is defined as G=(V,E) consisting of
 - Set V of vertices: V = {v₁, v₂, ..., vₙ}
 Number of vertices is denoted by |V| = n
 - Set E of unordered pairs $\{v_i, v_j\}$ termed edges
 - > Edges connect the vertices
- Maximum number of edges in an undirected graph is O(|V|²)

$$|E| \le {|V| \choose 2} = \frac{|V|(|V|-1)}{2} = O(|V|^2)$$

- Assumption: A vertex is never adjacent to itself
- For example, $\{v_1, v_1\}$ will not define an edge
- Many data structures can implement abstract undirected graphs
 - Adjacency matrices, Adjacency lists (as discussed later in the lecture)

Degree

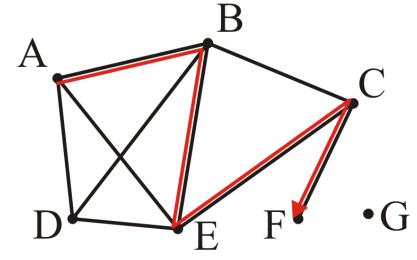
- Degree of a vertex is defined as the number of adjacent vertices
 - degree(A) = degree(D) = degree(C) = 3
 - degree(B) = degree(E) = 4
 - degree(F) = 1
 - degree(G) = 0



Vertices adjacent to a given vertex are its neighbors

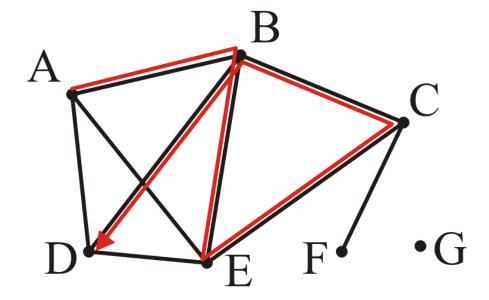
Path

- Path in an undirected graph is an ordered sequence of vertices
 - Consecutive vertices are connected through edges
- Path from vertex 0 to vertex k is $(v_0, v_1, v_2, \ldots, v_k)$
 - where $\{v_j 1, v_j\}$ is an edge for $j = 1, \ldots, k$
- Length of a path is equal to the number of edges
- Example: Path from A to F
 - Path: (A, B, E, C, F)
 - Length of the path is 4



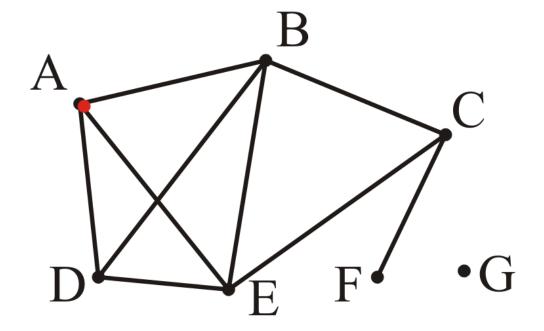
Path – Example

- Path of length 5: (A, B, E, C, B, D)
 - Repitition of vertex B



Path – Example

• A trivial path of length 0: (A)

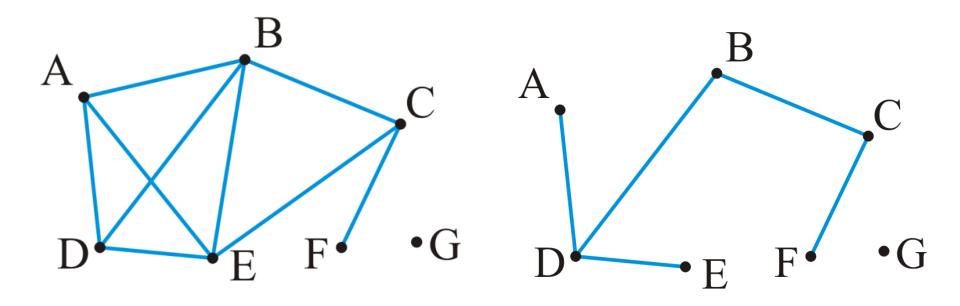


Path – Types

- A simple path has no repetitions other than perhaps the first and last vertices
- A cycle is a simple path of at least two vertices with the first and last vertices equal
 - Note: these definitions are not universal
- A loop is an edge from a vertex onto itself

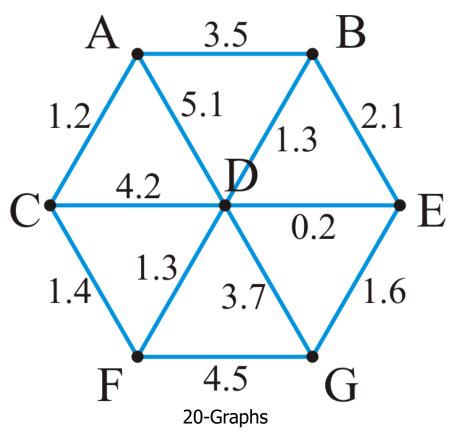
Subgraph

- A sub-graph of a graph G is defined by
 - Subset of the vertices
 - Subset of the edges that connected the subset of vertices in the original graph
- Every graph is a subgraph of itself



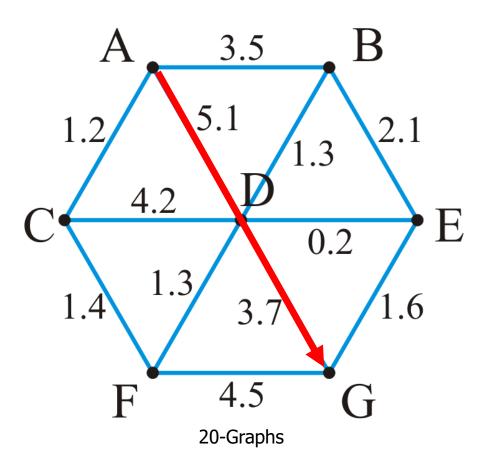
Weighted Graphs

- A weight may be associated with each edge in a graph
 - This could represent distance, energy consumption, cost, etc.
 - Such a graph is called a weighted graph
- Pictorially, we will represent weights by numbers next to the edges



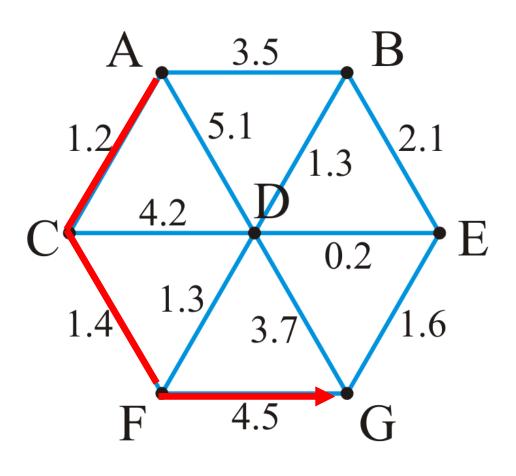
Weighted Graphs

- Length of a path within a weighted graph is the sum of all of the edges which make up the path
- The length of the path (A, D, G) in the following graph is 8.8



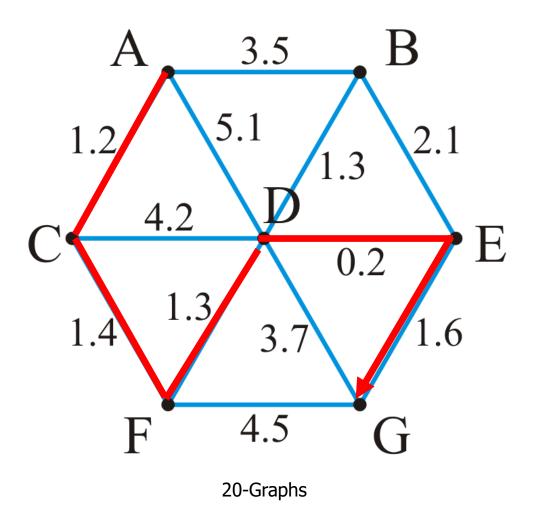
Weighted Graphs – Example

- Different paths may have different weights
 - Another path is (A, C, F, G) with length 1.2 + 1.4 + 4.5 = 7.1



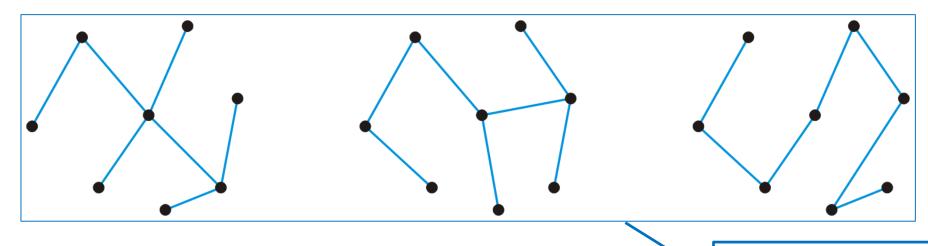
Weighted Graphs – Example

- Find the shortest path between two vertices A and G
- Shortest path is (A, C, F, D, E, G) with length 5.7



Trees

- A graph is a tree if it satisfies the following two conditions
 - Graph is connected
 - There is a unique path between any two vertices



- Consequences
 - The number of edges is |E| = |V| 1

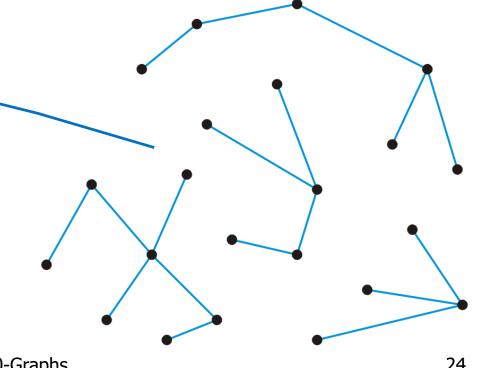
Three trees on same 8 vertices

- The graph is acyclic, that is, it does not contain any cycles
- Adding one more edge must create a cycle
- Removing any one edge creates two disjoint non-empty sub-graphs

Forest

- A forest is any graph that has no cycles
- Consequences
 - The number of edges is |E| < |V|
 - The number of trees is |V| |E|
 - Removing any one edge adds one more tree to the forest

- Forest with 22 vertices and 18 edges
- Four trees

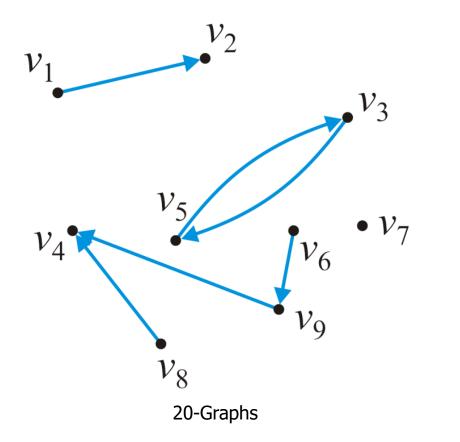


Directed Graphs

- In a directed graph, the edges on a graph are associated with a direction
 - Edges are ordered pairs (v_j, v_k) denoting a connection from v_j to v_k
 - The edge (v_j, v_k) is different from the edge (v_k, v_j)
- Streets are undirected graphs
 - In most cases, you can go two ways unless it is a one-way street

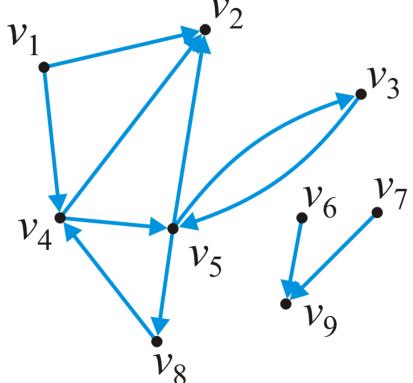
Directed Graphs

- Given our graph of nine vertices $V = \{v_1, v_2, \dots v_9\}$
 - These six pairs (v_i, v_k) are directed edges
 - E = { (v_1, v_2) , (v_3, v_5) , (v_5, v_3) , (v_6, v_9) , (v_8, v_4) , (v_9, v_4) }



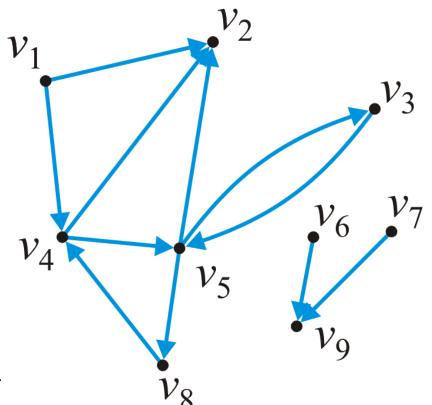
In and Out Degree

- Degree of a vertex must be modified to consider both cases:
 - Out-degree of a vertex is the number of vertices which are adjacent to the given vertex
 - Number of outgoing edges
 - In-degree of a vertex is the number of vertices which this vertex is adjacent to
 - > Number of incoming edges
- In this graph:
 - In-degree(v_1) = 0 out-degree(v_1) = 2
 - In-degree(v_5) = 2 out-degree(v_5) = 3



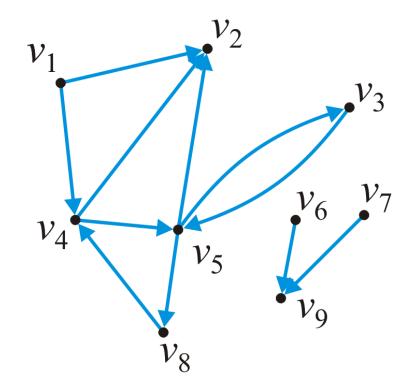
Path

- A path in a directed graph is an ordered sequence of vertices
 - $(v_0, v_1, v_2, \ldots, v_k)$
 - where $(v_j 1, v_j)$ is an edge for j = 1, ..., k
- A path of length 5 in this graph is
 - $(v_1, v_4, v_5, v_3, v_5, v_2)$
- A simple cycle of length 3 is
 - $-(v_8, v_4, v_5, v_8)$



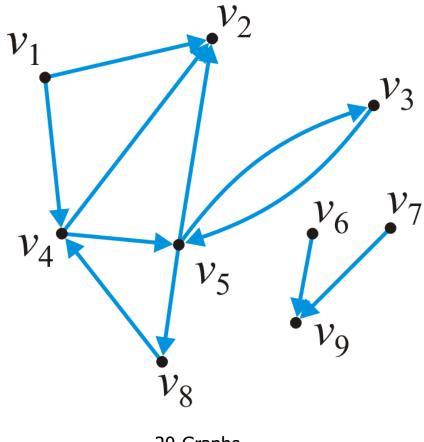
Connectedness

- Two vertices $v_{\tt j}$, $v_{\tt k}$ are said to be connected if there exists a path from $v_{\tt j}$ to $v_{\tt k}$
 - A graph is strongly connected if there exists a directed path between any two vertices
 - A graph is weakly connected if there exists a path between any two vertices that ignores the direction



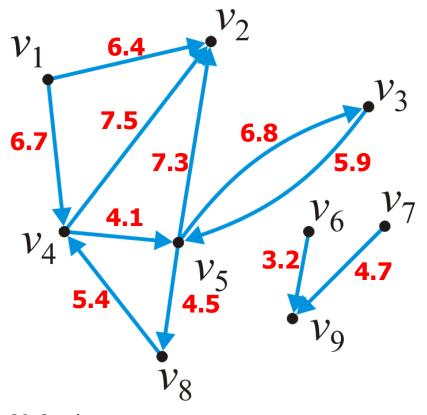
Connectedness – Example

- The sub-graph $\{v_3, v_4, v_5, v_8\}$ is strongly connected
- The sub-graph $\{v_1, v_2, v_3, v_4, v_5, v_8\}$ is weakly connected



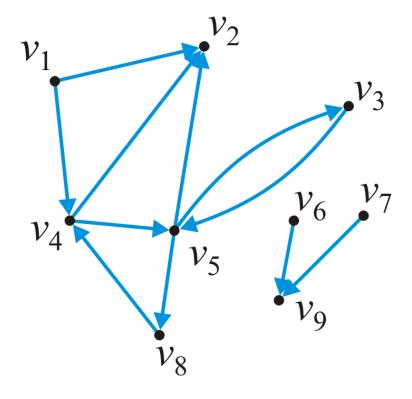
Weighted Directed Graphs

- Each edge is associated with a value
- If both (v_i, v_k) and (v_i, v_k) are edges
 - It is not required that they have the same weight



Representation

- How do we store the adjacency relations?
 - Adjacency matrix
 - Adjacency list

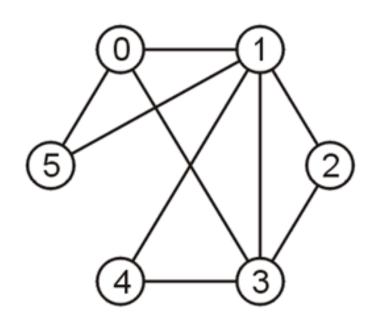


Adjacency Matrix

- Two dimensional matrix of size n x n where n = |V|
- a[i, j] = 0 (F) if there is no edge between vertices v_i and v_j
- a[i, j] = 1 (T) if there is an edge between vertices v_i and v_j
- Adjacency matrix of undirected graphs is symmetric

$$- a[i, j] = a[j, i]$$

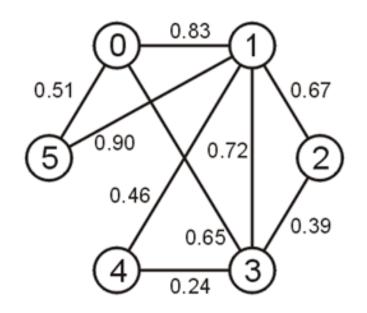
	0	1	2	3	4	5
0	F	Т	F	Т	F	Т
1	Т	F	Т	Т	Т	Т
2	F	Т	F	Т	F	F
3	Т	Т	Т	F	Т	F
4	F	Т	F	Т	F	F
5	Т	Т	F	F	F	F



Adjacency Matrix – Weighted Graph

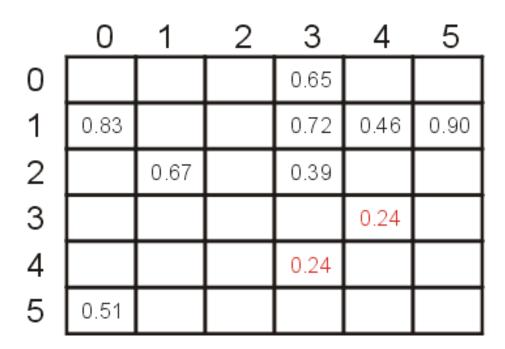
- The matrix entry [j, k] is set to the weight of the edge (v_i, v_k)
- How to indicate absence of an edge in the graph?

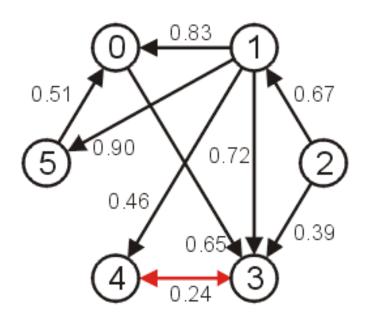
	0	1	2	3	4	5
0		0.83		0.65		0.51
1	0.83		0.67	0.72	0.46	0.90
2		0.67		0.39		
3	0.65	0.72	0.39		0.24	
4		0.46		0.24		
5	0.51	0.90				



Adjacency Matrix – Directed Graph

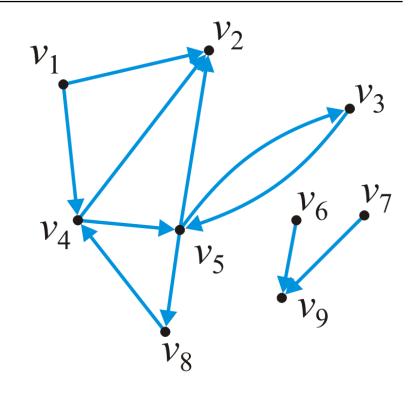
• For directed graph the matrix would not necessarily be symmetric





Adjacency Matrix – Analysis

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

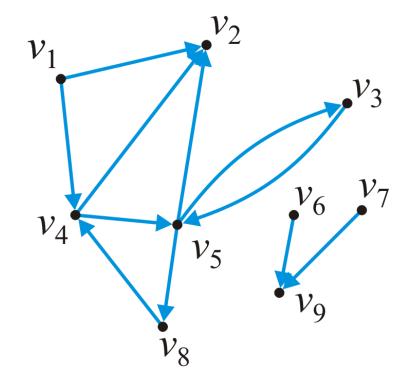


- Requires memory: 0(|V|²)
- Determining if v_i is adjacent to v_k: 0(1)
- Finding all neighbors of v_i: O(|V|)

Adjacency Matrix – Problem

- Very sparsely populated
 - Out of 81 cells only 11 are 1 (or T)

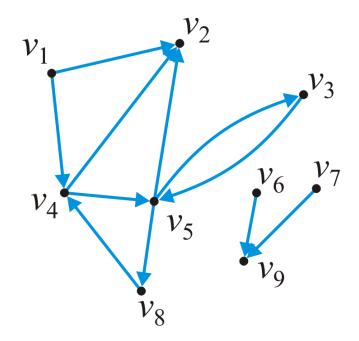
	1	2	3	4	5	6	7	8	9
1		1		1					
2									
3					1				
4		1			1				
5		1	1					1	
6									1
7									1
8				1					
9									



Adjacency List

- Each vertex is associated with a list of its neighbors
 - A vertex w is inserted in the list for vertex v if edge (v, w) exists

$$\begin{array}{cccc}
1 & \bullet \rightarrow 2 \rightarrow 4 \\
2 & \bullet \\
3 & \bullet \rightarrow 5 \\
4 & \bullet \rightarrow 2 \rightarrow 5 \\
5 & \bullet \rightarrow 2 \rightarrow 3 \rightarrow 8 \\
6 & \bullet \rightarrow 9 \\
7 & \bullet \rightarrow 9 \\
8 & \bullet \rightarrow 4 \\
9 & \bullet
\end{array}$$



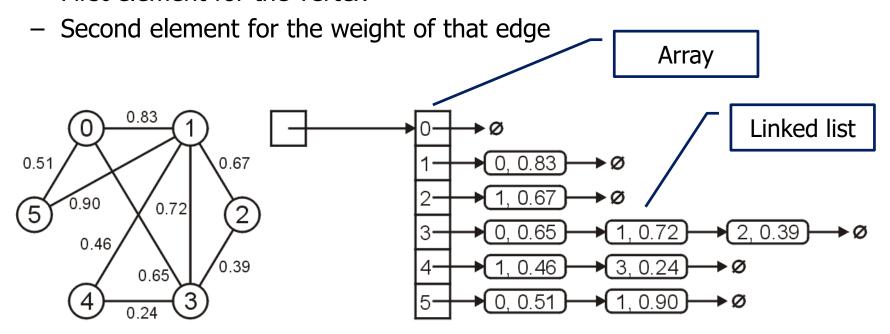
Requires memory: O(|V| + |E|)

20-Graphs

38

Adjacency List – Weighted Graphs

- An adjacency list for a weighted graph contains two elements
 - First element for the vertex



- When the vertices are identified by a name (i.e., string)
 - Hash-table of lists is used to implement the adjacency list

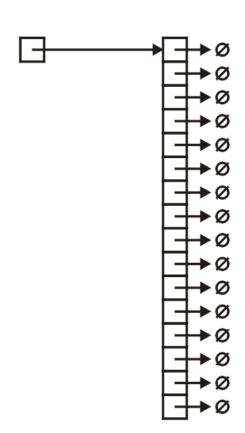
Adjacency List – Implementation

Node to store adjacent vertex and weight of the edge

```
class SingleNode {
   private:
      int adjacent_vertex;
      double edge_weight;
      SingleNode * next_node;
};
```

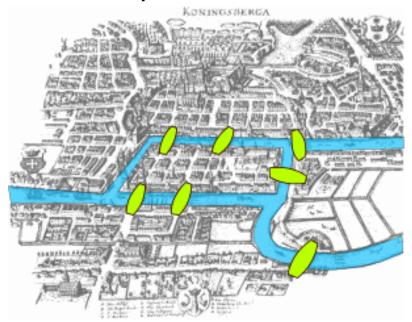
Define and create Array

```
SingleNode * array;
array = new SingleNode[16];
```



Graph Problems – Euler Tour

- A sequence of vertices that traverse all edges in the graph exactly once
 - Leonhard Euler in 1736
 - Laid the foundations of graph theory
- Problem: To devise a walk through the city that would cross each of the seven bridges of Königsberg once and only once
 - Euler proved problem has no solution



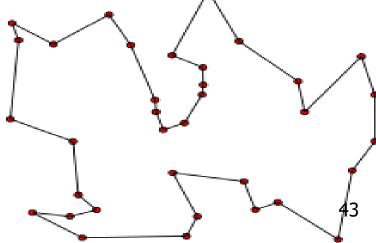
Graph Problems – Traveling Salesman

- A salesman wishes to
 - Visit a number of towns, and then
 - Return to his starting town
- Given the travelling times between towns, how should the travel be planned, so that:
 - He visits each town exactly once, and
 - He travels in as short time as possible

 Problem: Given a weighted graph G, provide shortest cycle that contains all vertices in G

20-Graphs

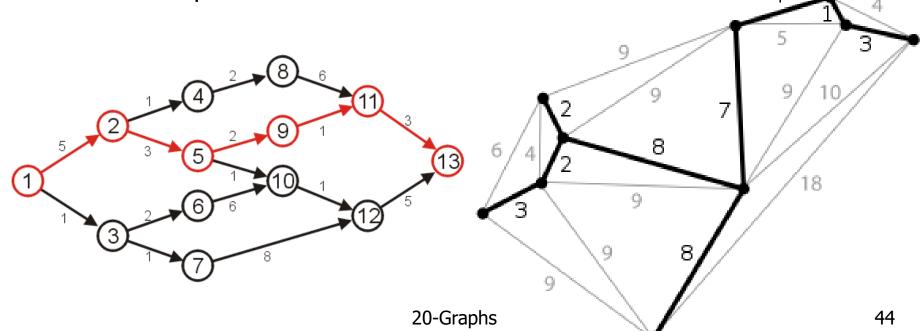
NP-Hard problem



Graph Problems – Others

- Minimum-cost spanning tree
 - Given a weighted graph G, determine a spanning tree with minimum total edge cost
- Single-source shortest path

Given a weighted graph G and a source vertex v in G, determine the shortest paths from v to all other vertices in G



Graph Traversal

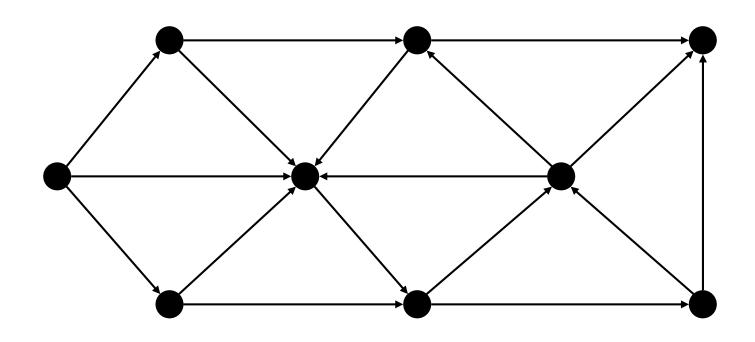
- Given a graph G = (V, E), directed or undirected
 - Goal is to methodically explore every vertex and every edge
- Traversals of graphs are also called searches
- We can use either breadth-first or depth-first traversals
 - Breadth-first requires a queue
 - Depth-first requires a stack

Breadth-First Search

- Choose any vertex, mark it as visited and enqueue it onto queue
- While the queue is not empty
 - Dequeue top vertex v from the queue
 - For each vertex adjacent to v that has not been visited
 - > Mark it visited, and
 - > Enqueue it onto the queue

```
1:create a queue Q
2:mark v as visited and put v into Q
3:while Q is non-empty
4: remove the head u of Q (Dequeue)
5: mark and enqueue all (unvisited) neighbors of u
```

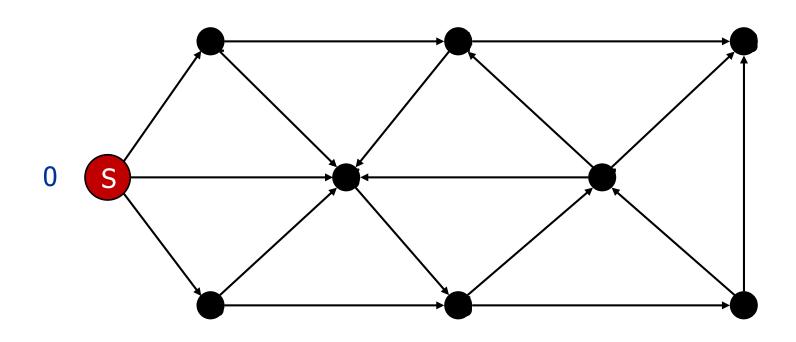
- The above algorithm continues until the queue is empty!
 - If there are no unvisited vertices, the graph is connected



Undiscovered
Discovered
Top of queue
Finished

Queue (Q):		

1: Create a Queue Q

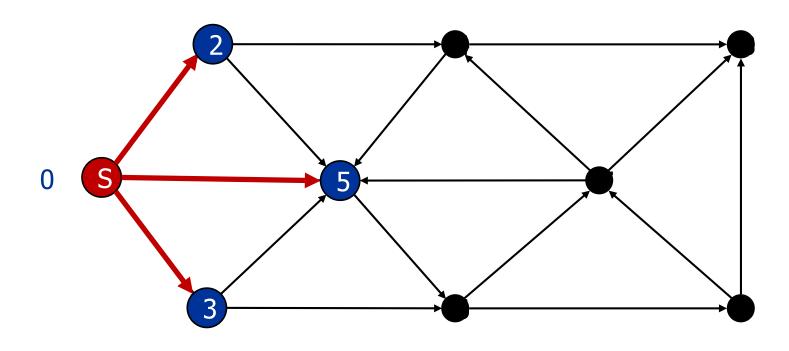


Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

S

2: Mark S as visited and put S into Q



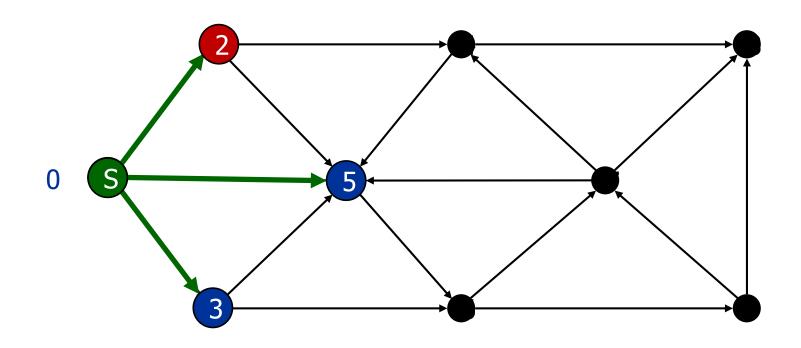
Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

S

```
3: While Q not empty
```

4:
$$v = dequeue Q (i.e., S)$$



Undiscovered
Discovered
Top of queue
Finished

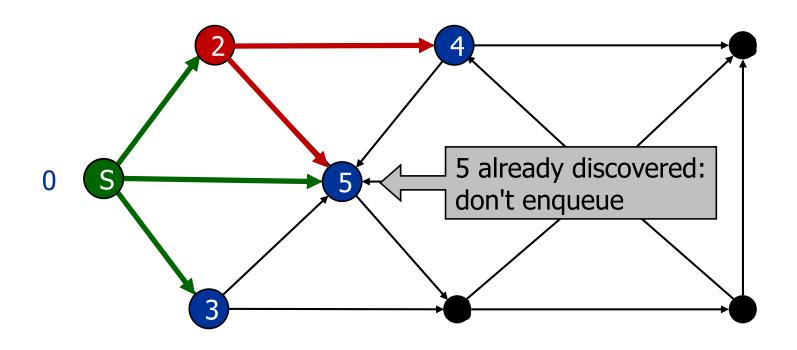
Queue (Q):

235

```
3: While Q not empty
4: v = dequeue Q (i.e., 2)
```

5: mark & enqueue all (unvisited) neighbors of v

21-Graph Traversal



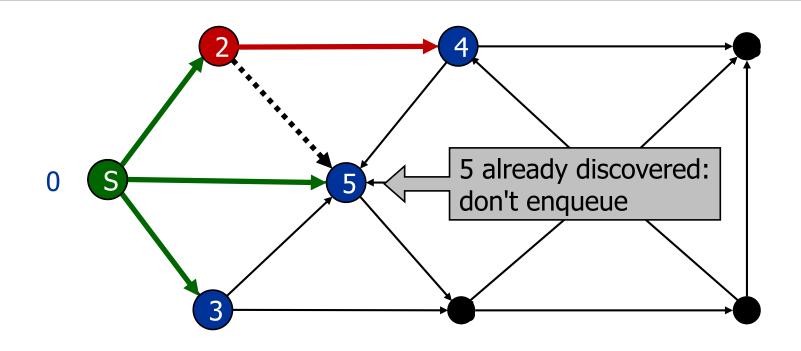
Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

235

```
3: While Q not empty
```

4: v = dequeue Q (i.e., 2)



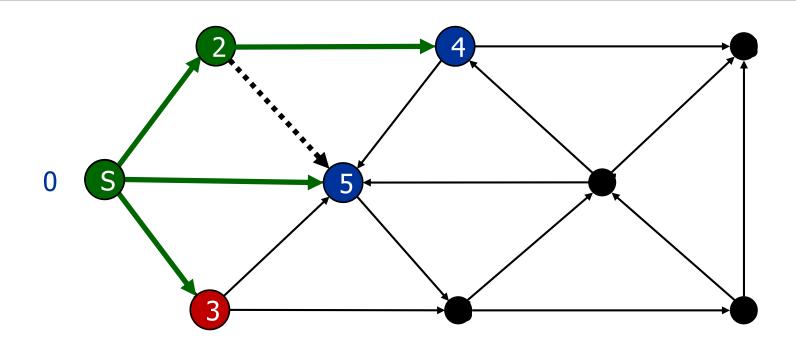
Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

235

```
3: While Q not empty
```

4:
$$v = dequeue Q (i.e., 2)$$

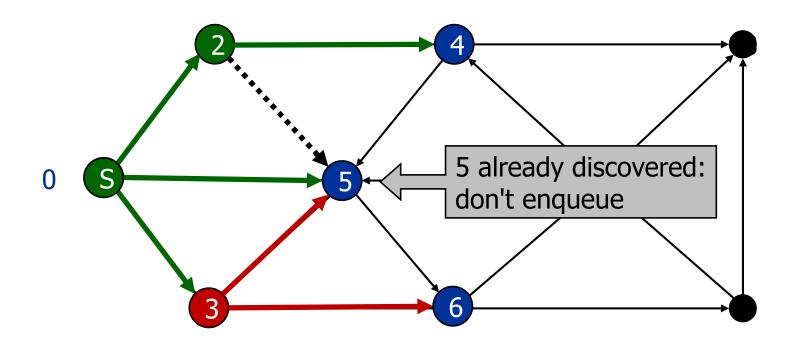


Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

354

```
3: While Q not empty
4: v = dequeue Q (i.e., 3)
5: mark & enqueue all (unvisited) neighbors of v
```



Undiscovered
Discovered
Top of queue
Finished

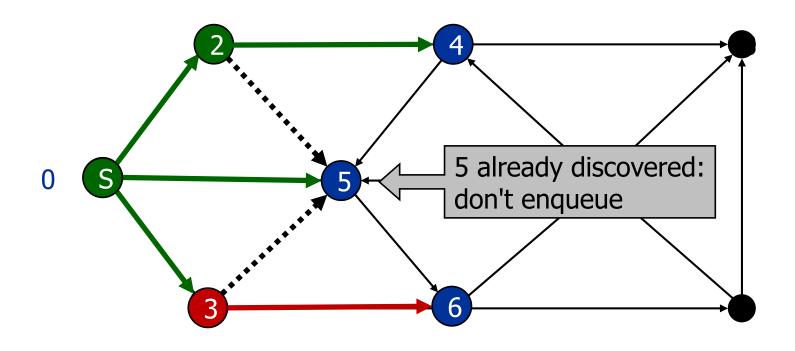
Queue (Q):

354

```
3: While Q not empty
4: v = dequeue Q (i.e., 3)
```

5: mark & enqueue all (unvisited) neighbors of v

21-Graph Traversal

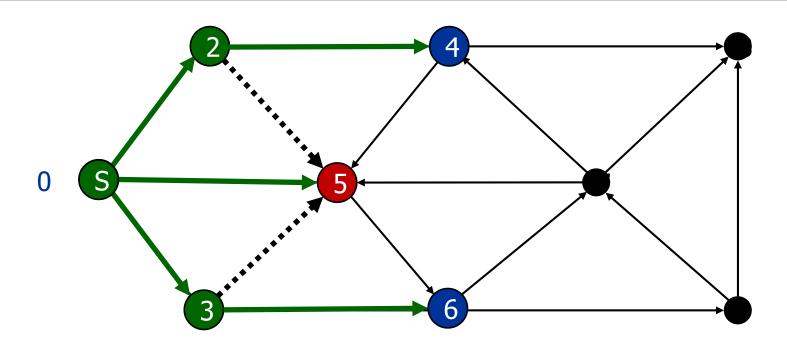


Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

354

```
3: While Q not empty
4: v = dequeue Q (i.e., 3)
5: mark & enqueue all (unvisited) neighbors of v
```



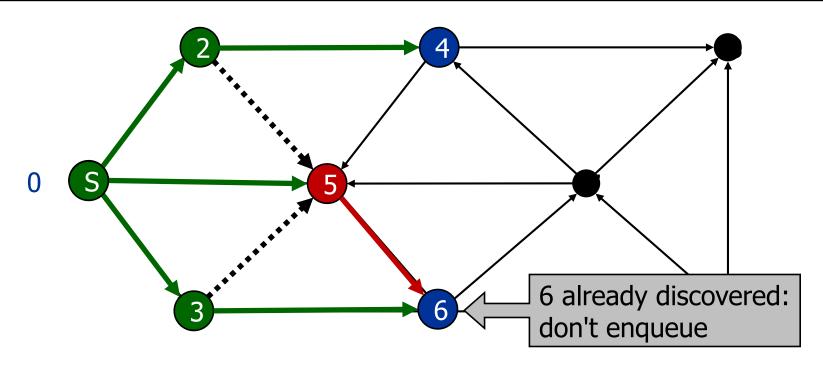
Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

5 4 6

```
3: While Q not empty
```

4:
$$v = dequeue Q (i.e., 5)$$



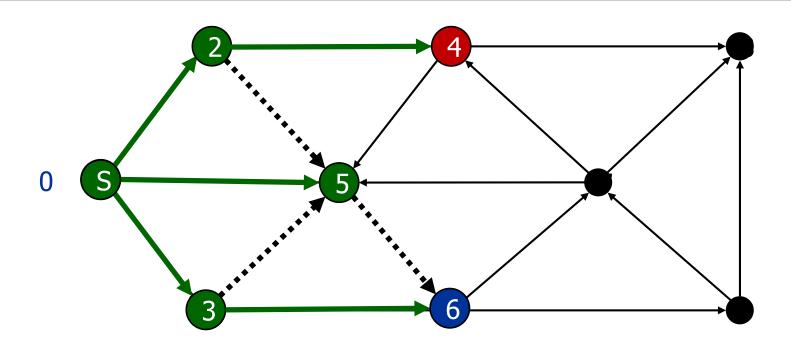
Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

5 4 6

```
3: While Q not empty
```

4: v = dequeue Q (i.e., 5)



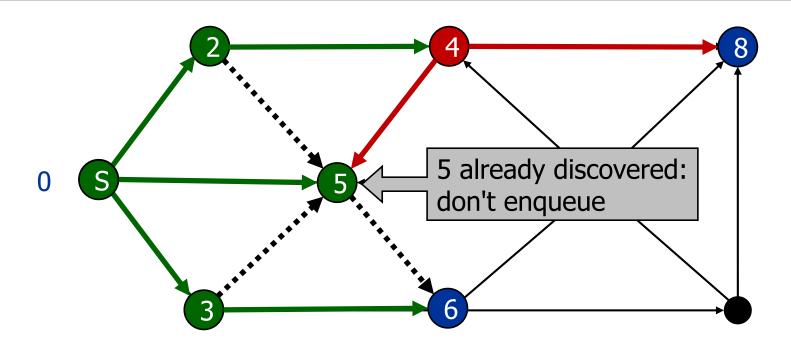
Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

46

```
3: While Q not empty
```

4: v = dequeue Q (i.e., 4)



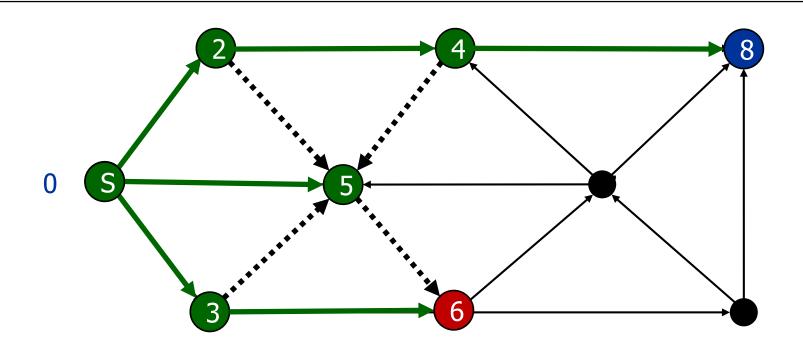
Undiscovered
Discovered
Top of queue
Finished

```
Queue (Q):
```

46

```
3: While Q not empty
```

4: v = dequeue Q (i.e., 4)



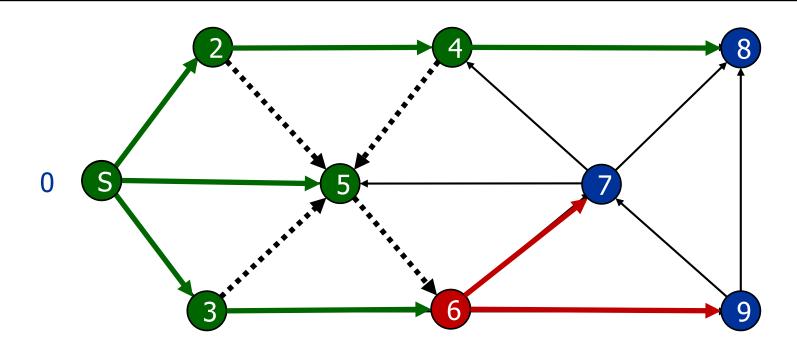
Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

68

```
3: While Q not empty
```

4: v = dequeue Q (i.e., 6)



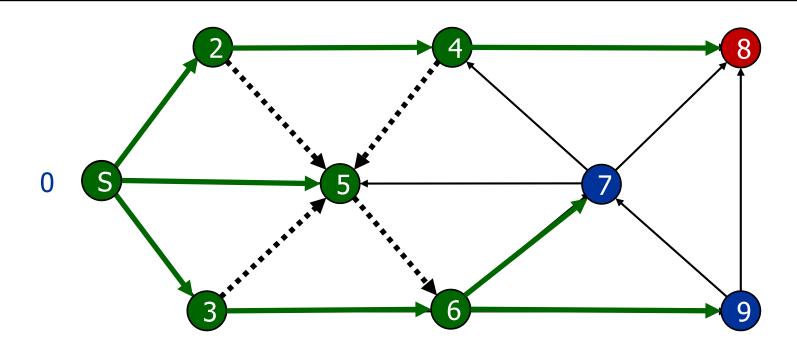
Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

68

```
3: While Q not empty
```

4: v = dequeue Q (i.e., 6)



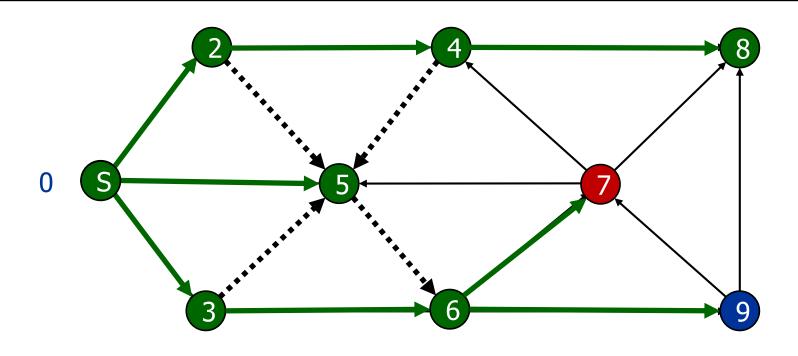
Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

879

```
3: While Q not empty
```

4: v = dequeue Q (i.e., 8)



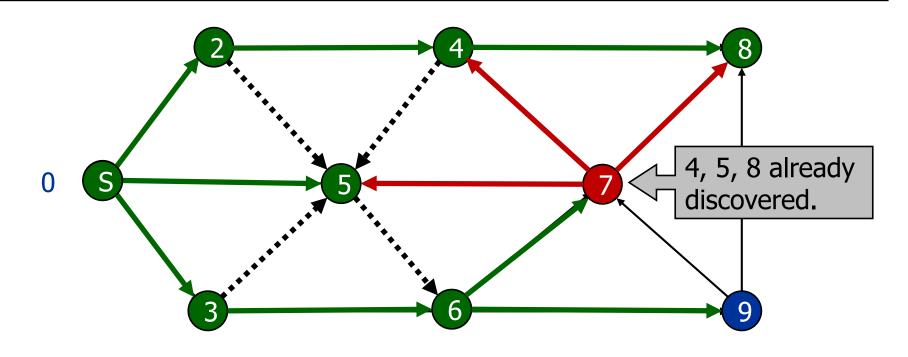
Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

7 9

```
3: While Q not empty
```

4: v = dequeue Q (i.e., 7)



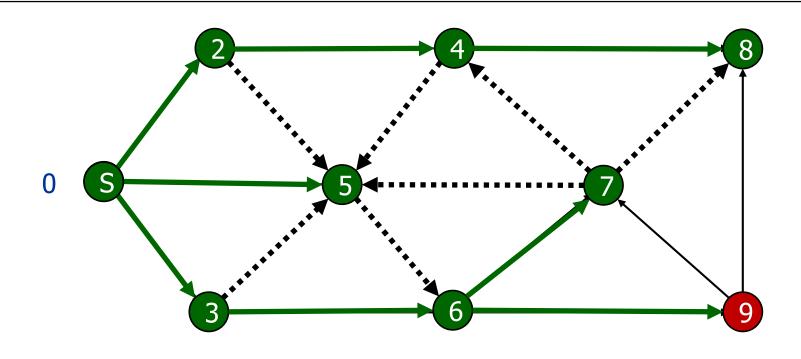
Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

7 9

```
3: While Q not empty
```

4: v = dequeue Q (i.e., 7)



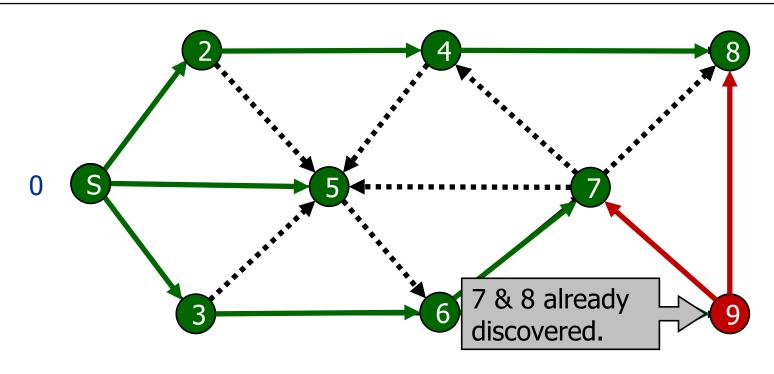
Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

9

3: While Q not empty

4: v = dequeue Q (i.e., 9)



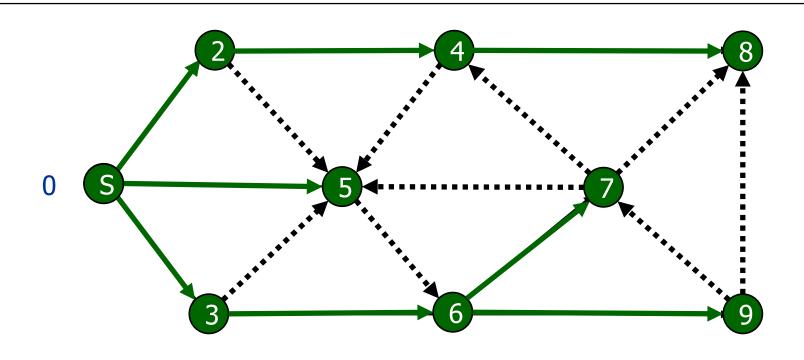
Undiscovered
Discovered
Top of queue
Finished

```
Queue (Q):
```

9

```
3: While Q not empty
```

4:
$$v = dequeue Q (i.e., 9)$$

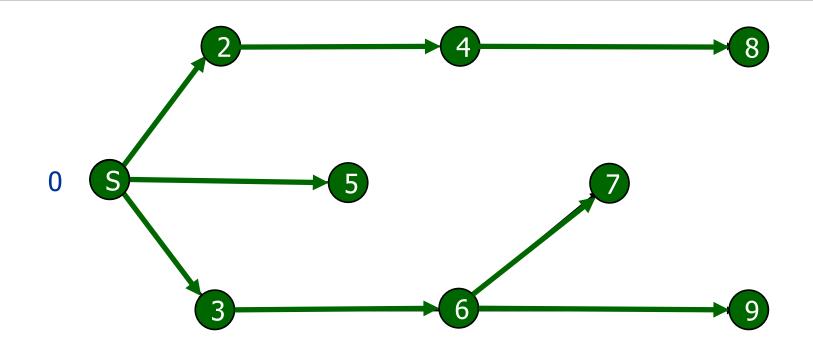


Undiscovered
Discovered
Top of queue
Finished

Queue (Q):

3: While Q not empty

4: v = dequeue Q (i.e., NULL)



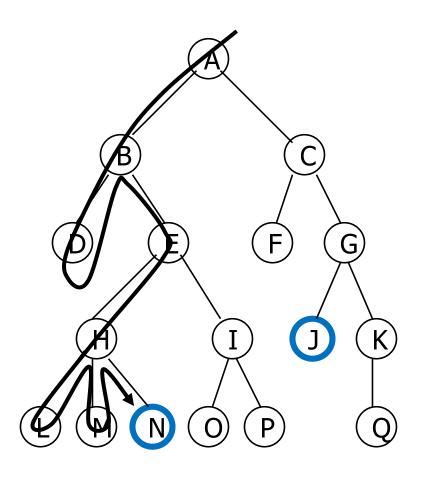
Breadth-First Search (BFS) tree rooted at S containing all nodes of the graph

Breadth-First Search – Properties

- Given a graph G=(V,E) and source vertex S, the following holds for the BFS algorithm
 - Systematically explores the edges of G to "discover" every vertex reachable from S
 - Creates a BFS tree rooted at S that contains all such vertices
 - Discovers all vertices at distance k from S before discovering any vertices at distance k+1

Depth-First Search – Trees

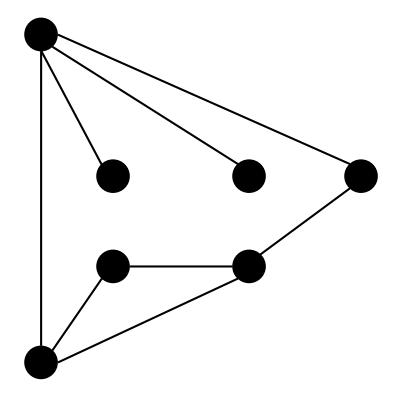
- A depth-first search (DFS) explores a path all the way to a leaf before backtracking and exploring another path
- For example, after searching A, then B, then D, the search backtracks and tries another path from B
- N will be found before J
- Node are explored in the order A B D E H L M N I O P C F G J K Q



Depth-First Search

- Choose any vertex, mark it as visited
- From that vertex:
 - If there is another adjacent vertex not yet visited, go to it
 - Otherwise, go back to the most previous vertex that has not yet had all of its adjacent vertices visited and continue from there
- Continue until no visited vertices have unvisited adjacent vertices

```
Create a stack S
Mark v as visited and push v onto S
while S is non-empty
peek at the top u of S
if u has an (unvisited) neighbor w
mark w and push it onto S
else
pop S
```



Adjacency List

A: FCBG

B: A

C: A

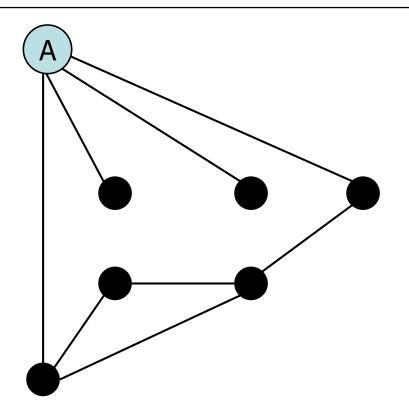
D: FE

E: GFD

F: AED

G: EA





Adjacency List

A: FCBG

B: A

C: A

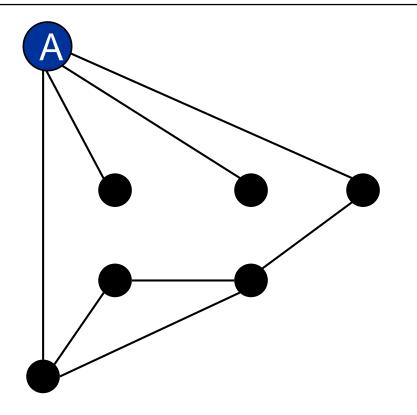
D: FE

E: GFD

F: AED

G: EA





Adjacency List

A: FCBG

B: A

C: A

D: FE

E: GFD

F: AED

G: EA

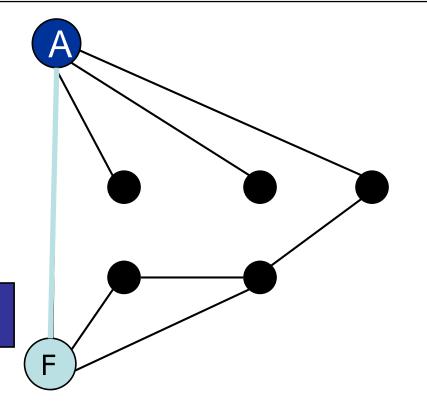
Undiscovered
Marked
Active
Finished

visit(A)

(A, F) (A, C) (A, B) (A, G)

Stack

21-Graph Traversal



Adjacency List

A: FCBG

B: A

C: A

D: FE

E: GFD

F: AED

G: EA

Undiscovered
Marked
Active
Finished

F newly

discovered

visit(A)
(A, F) (A, C) (A, B) (A, G)
Stack
21-Graph Traversal

Adjacency List

A: FCBG

B: A

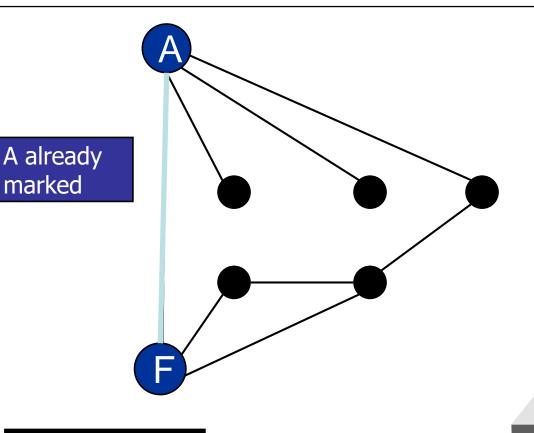
C: A

D: FE

E: GFD

F: AED

G: EA



Undiscovered

Marked

marked

Active

Finished

visit(F)

(F, A) (F, E) (F, D)

(A, F) (A, C) (A, B) (A, G)

Stack

21-Graph Traversal

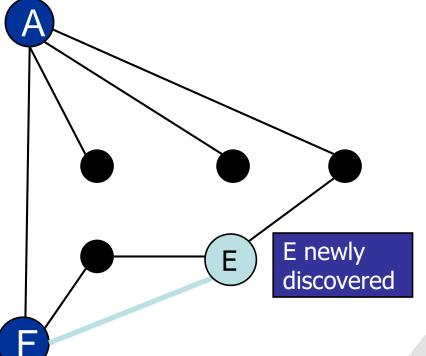
76

Undiscovered

Marked

Active

Finished



Adjacency List

A: FCBG

B: A

C: A

D: FE

E: GFD

F: AED

G: EA

visit(F)

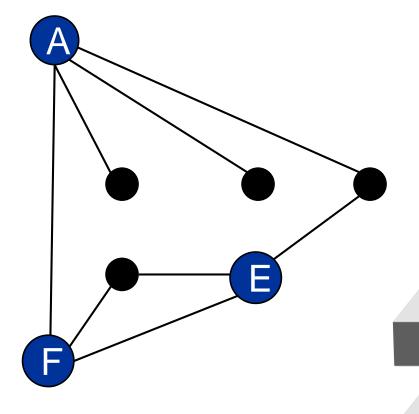
(F, A) (F, E) (F, D)



(A, F) (A, C) (A, B) (A, G)

Stack

21-Graph Traversal



Undiscovered

Marked

Active

Finished

Adjacency List

A: FCBG

B: A

C: A

D: FE

E: GFD

F: AED

G: EA

visit(E)

(E, G) (E, F) (E, D)

(F, A) (F, E) (F, D)

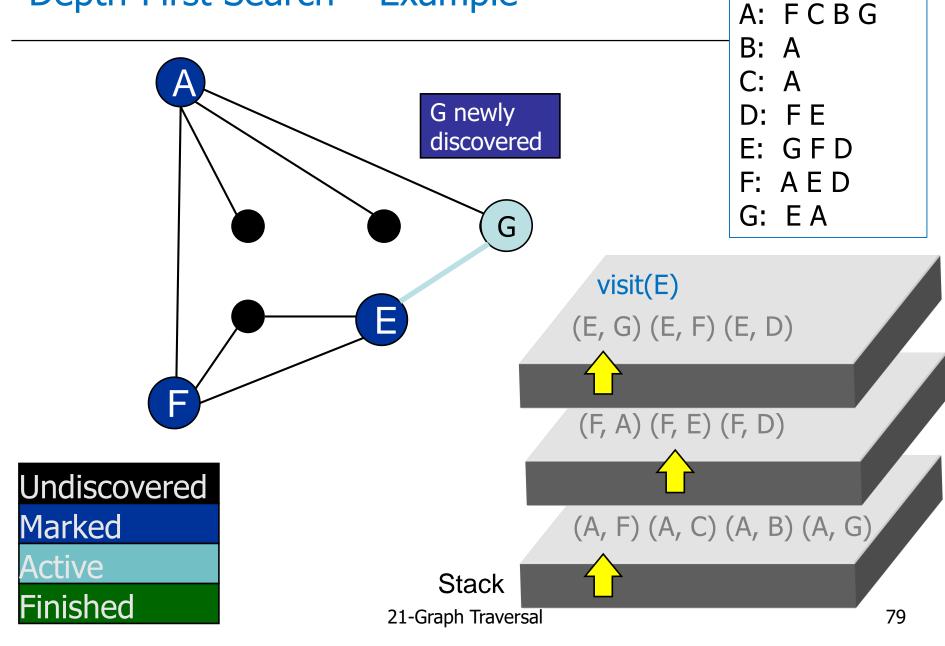


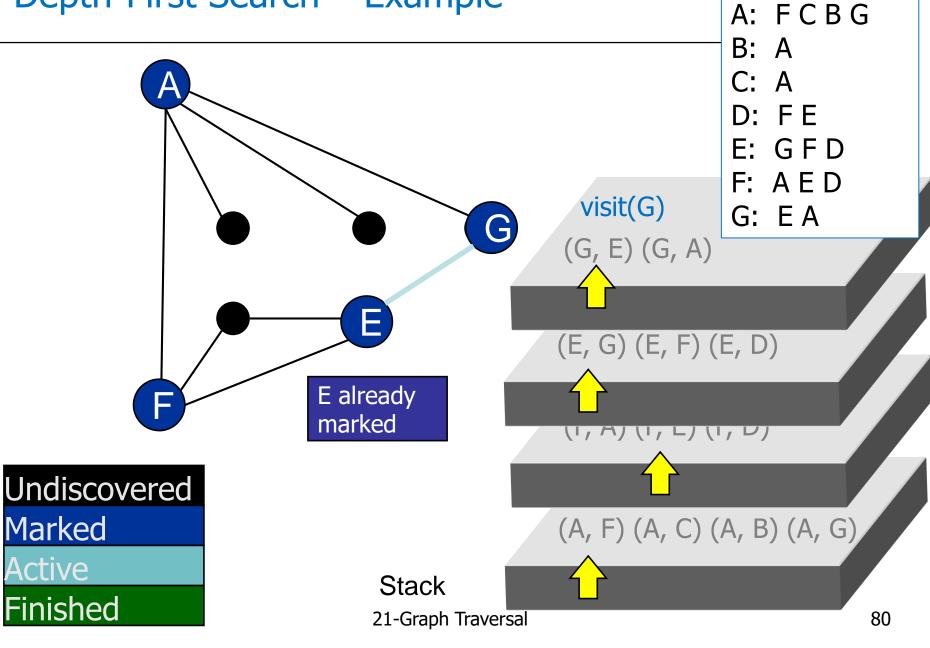
(A, F) (A, C) (A, B) (A, G)

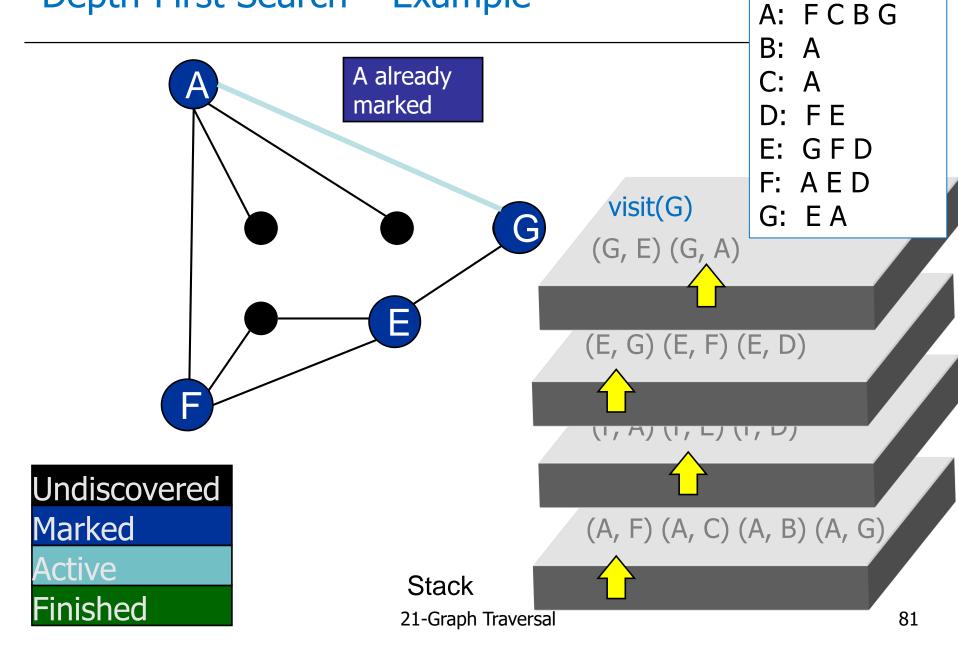
Stack

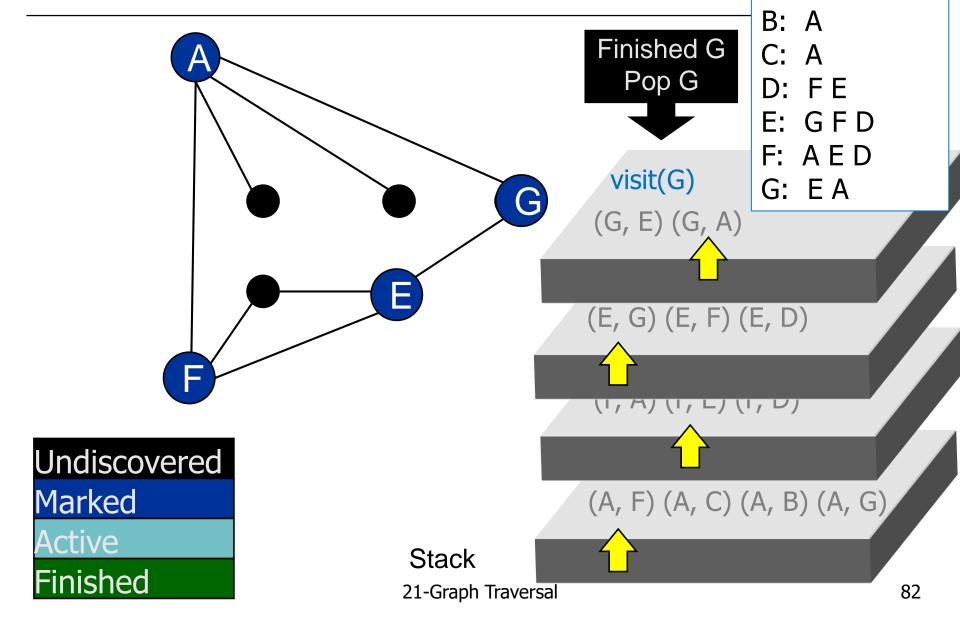
21-Graph Traversal

78

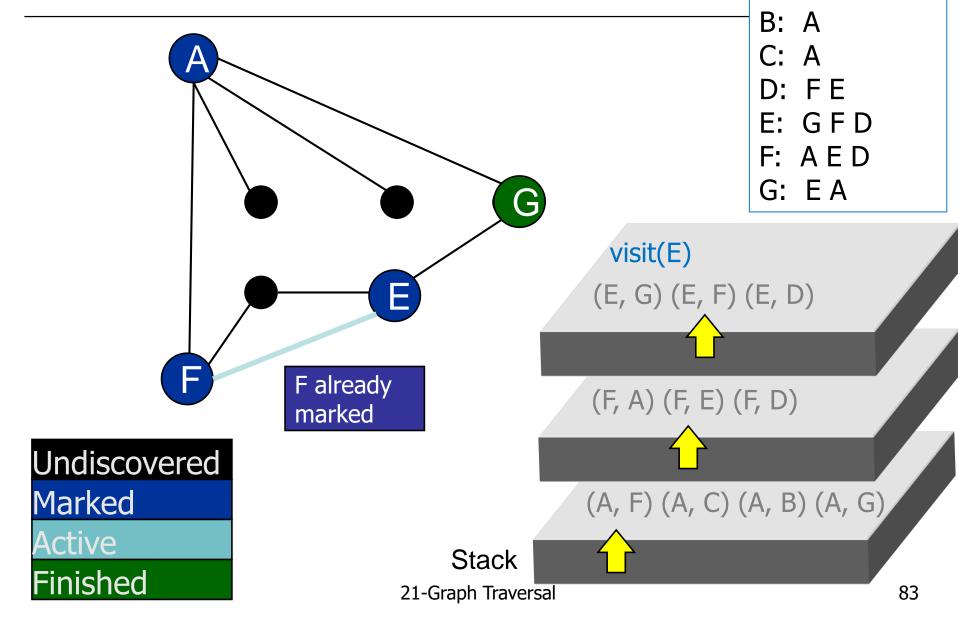




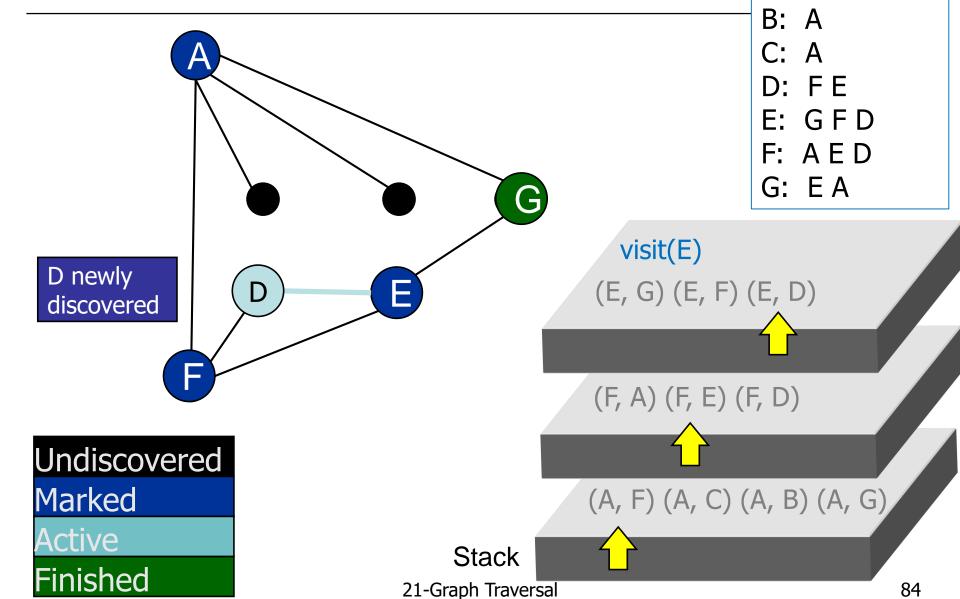




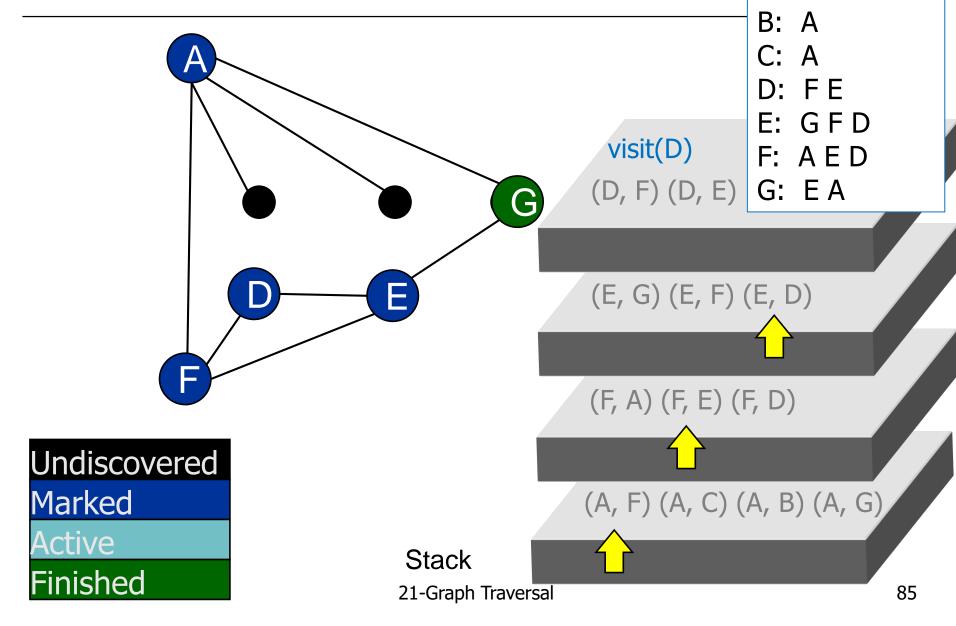
Adjacency List



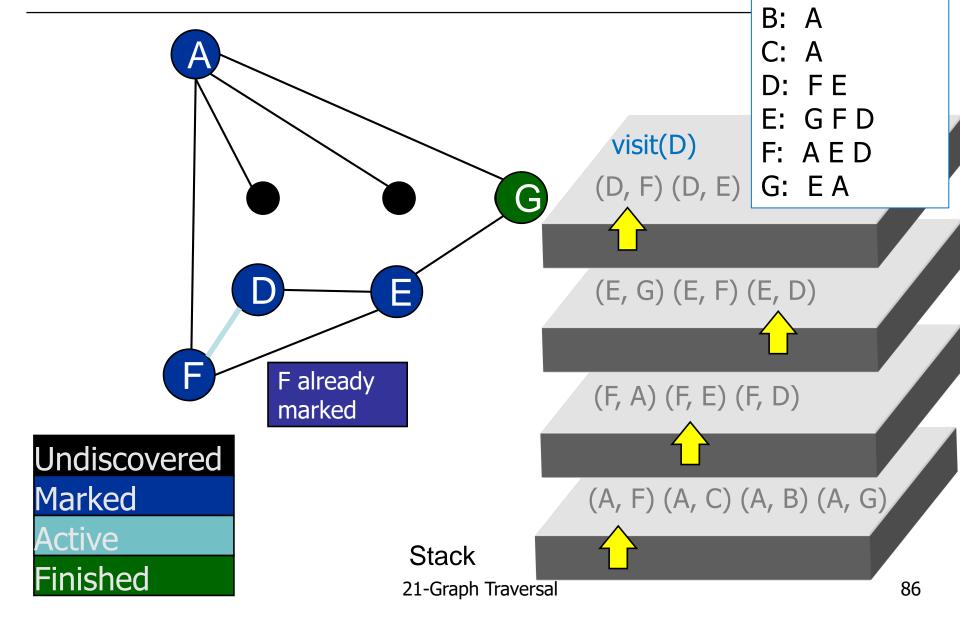
Adjacency List



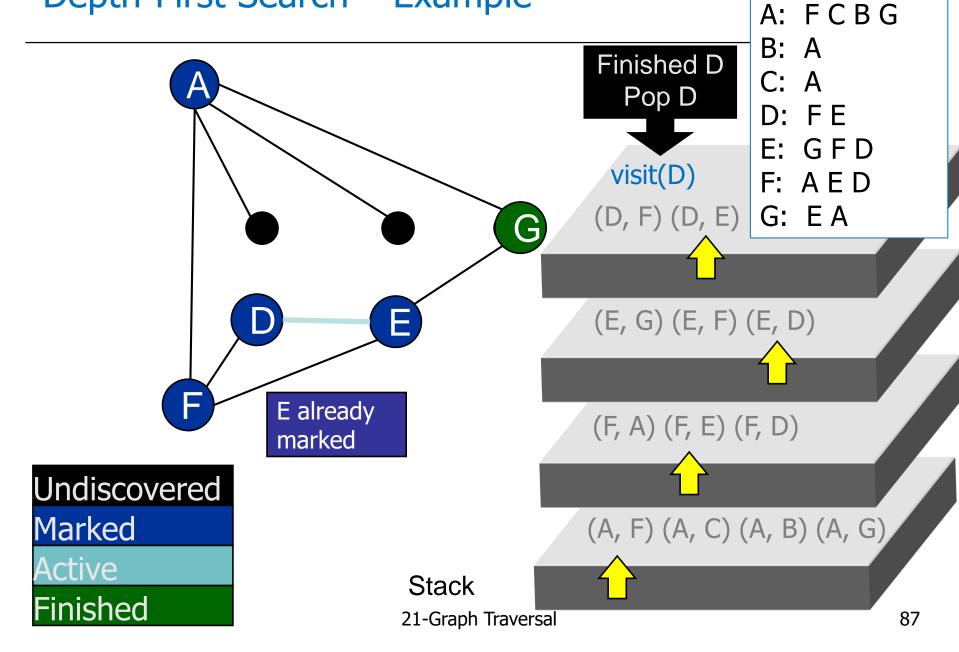
Adjacency List

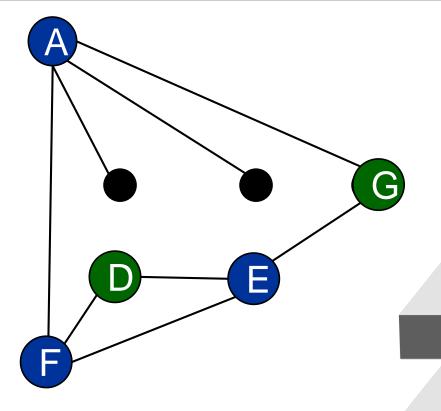


Adjacency List



Adjacency List





Undiscovered Marked

Active

Finished

Adjacency List

A: FCBG

B: A

C: A

D: FE

E: GFD

F: AED

G: EA

visit(E)

(E, G) (E, F) (E, D)



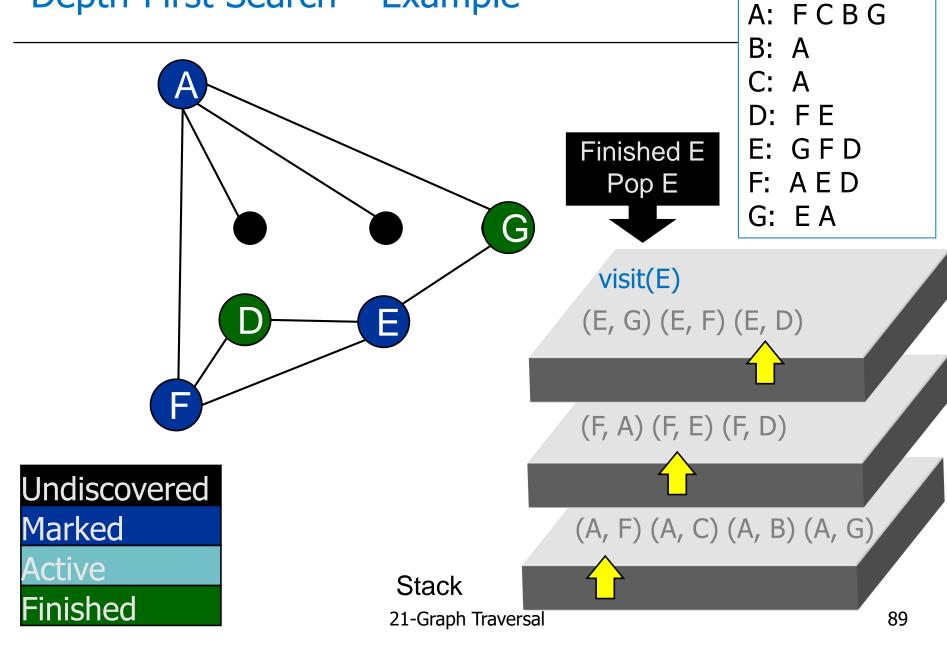
(F, A) (F, E) (F, D)



(A, F) (A, C) (A, B) (A, G)

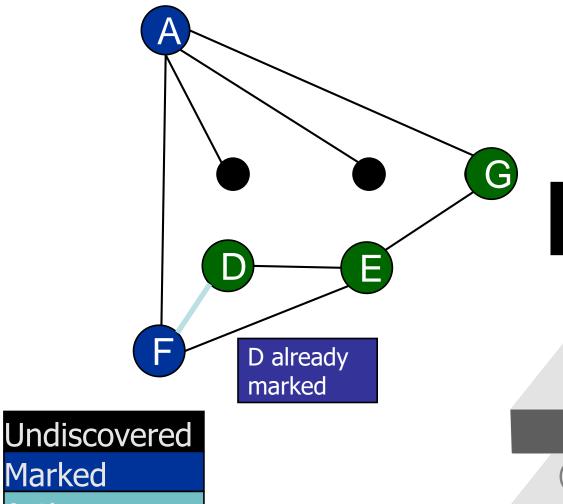


21-Graph Traversal



Active

Finished



Adjacency List

A: FCBG

B: A

C: A

D: FE

E: GFD

F: A E D

G: EA

visit(F)

Finished F

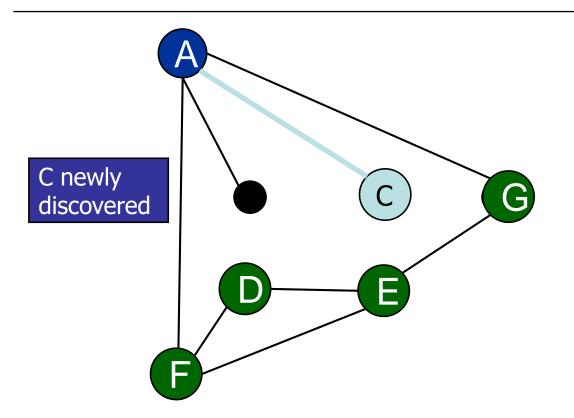
Pop F

(F, A) (F, E) (F, D)



Stack

21-Graph Traversal



Adjacency List

A: FCBG

B: A

C: A

D: FE

E: GFD

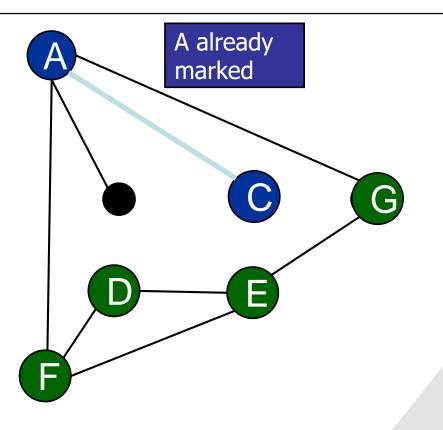
F: A E D

G: EA

Undiscovered
Marked
Active
Finished

Stack 21-Graph Traversal

visit(A)
(A, F) (A, C) (A, B) (A, G)



Adjacency List

A: FCBG

B: A

C: A

D: FE

E: GFD

F: AED

G: EA

Undiscovered

Marked

Active

Finished

visit(C)

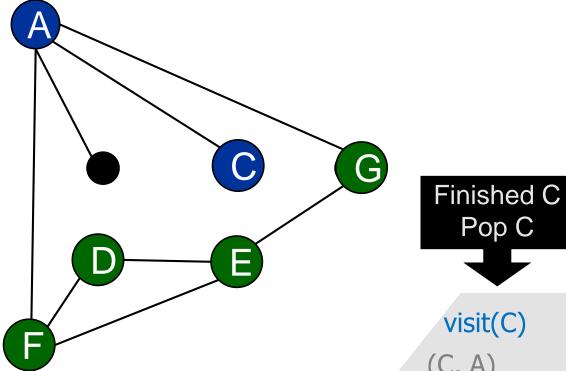
(C, A)



(A, F) (A, C) (A, B) (A, G)

Stack

21-Graph Traversal



Adjacency List

A: FCBG

B: A

C: A

D: FE

E: GFD

F: A E D

G: EA

visit(C)

Pop C

(C, A)



(A, F) (A, C) (A, B) (A, G)

Stack

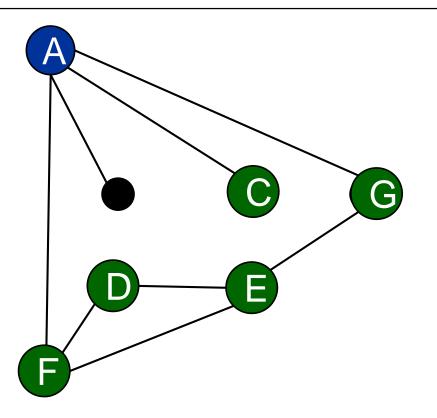
21-Graph Traversal

Active

Marked

Finished

Undiscovered



Adjacency List

A: FCBG

B: A

C: A

D: FE

E: GFD

F: AED

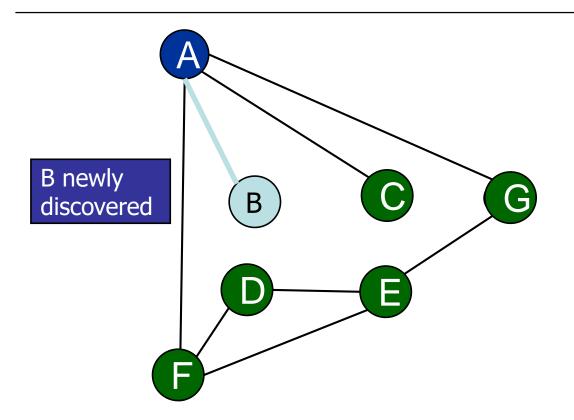
G: EA

Undiscovered
Marked
Active
Finished

Stack 21-Graph Traversal

(A, F) (A, C) (A, B) (A, G)

visit(A)



Adjacency List

A: FCBG

B: A

C: A

D: FE

E: GFD

F: AED

G: EA

Undiscovered
Marked
Active
Finished

Stack 21-Graph Traversal

(A, F) (A, C) (A, B) (A, G)

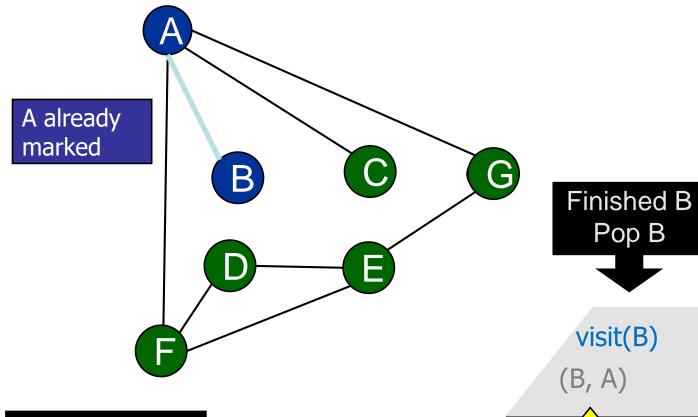
visit(A)

Undiscovered

Marked

Active

Finished



Adjacency List

A: FCBG

B: A

D: FE

E: GFD

F: A E D

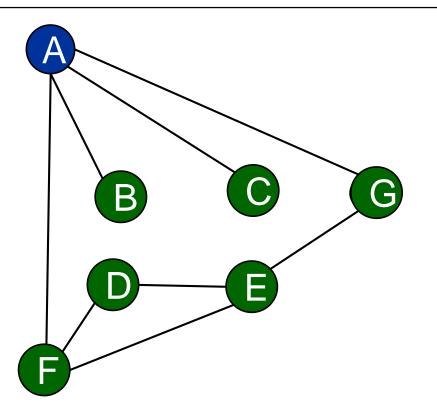
G: EA



(A, F) (A, C) (A, B) (A, G)

Stack

21-Graph Traversal



Adjacency List

A: FCBG

B: A

C: A

D: FE

E: GFD

F: AED

G: EA

Undiscovered

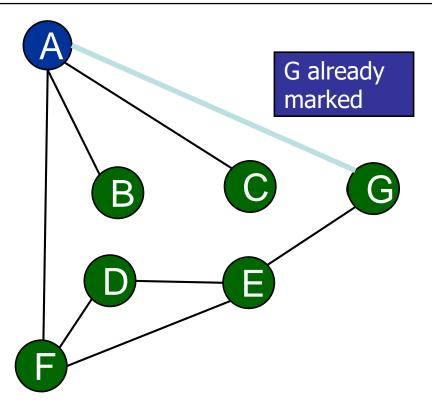
Marked

Active

Finished

Stack 21-Graph Traversal

visit(A)
(A, F) (A, C) (A, B) (A, G)



Adjacency List

A: FCBG

B: A

C: A

D: FE

E: GFD

F: A E D

G: EA

Finished A
Pop A

Undiscovered

Marked

Active

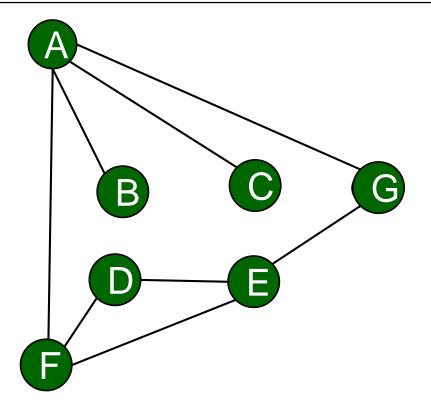
Finished

visit(A)

(A, F) (A, C) (A, B) (A, G)

Stack

21-Graph Traversal



Undiscovered
Marked
Active
Finished

A: FCBG

B: A

C: A

D: FE

E: GFD

F: AED

G: EA

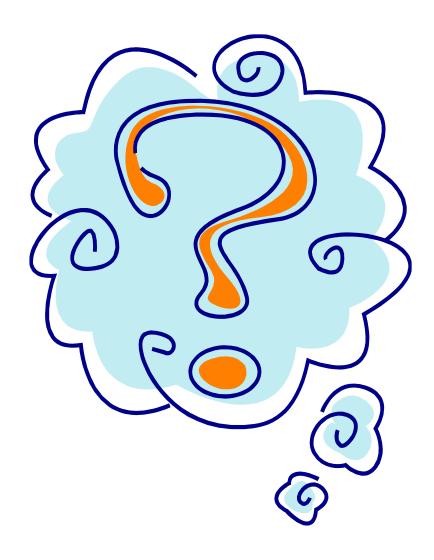
BFS vs. DFS

- Depending on the application, either DFS or BFS could be advantageous
- Example: Consider your family tree
 - If you are searching for some of your siblings cousins then it would be safe to assume that person would be on the bottom of the tree
 - Which approach is better in this case?
 - > In general, both approaches have the same time complexity
 - > In worst case, they need to visit all the nodes

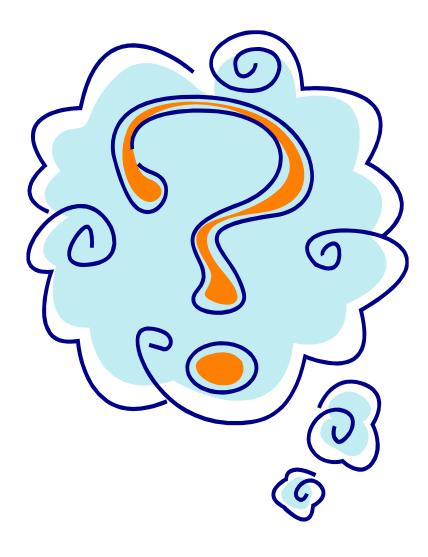
Applications of Graph Traversal

- Determining connectedness and finding connected sub-graphs
- Construct a BFS or DFS tree/forest from a graph
- Determining the path length from one vertex to all others
 - Find the shortest path from a vertex s to a vertex v (BFS)

Any Question So Far?



Any Question So Far?



20-Graphs