# Graphs

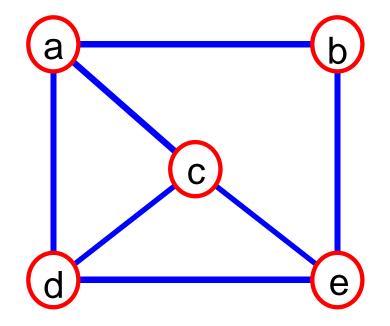
# What is a Graph?

A graph G = (V,E) is composed of:

V: a finite, nonempty set of vertices

E: set of edges connecting the vertices in V

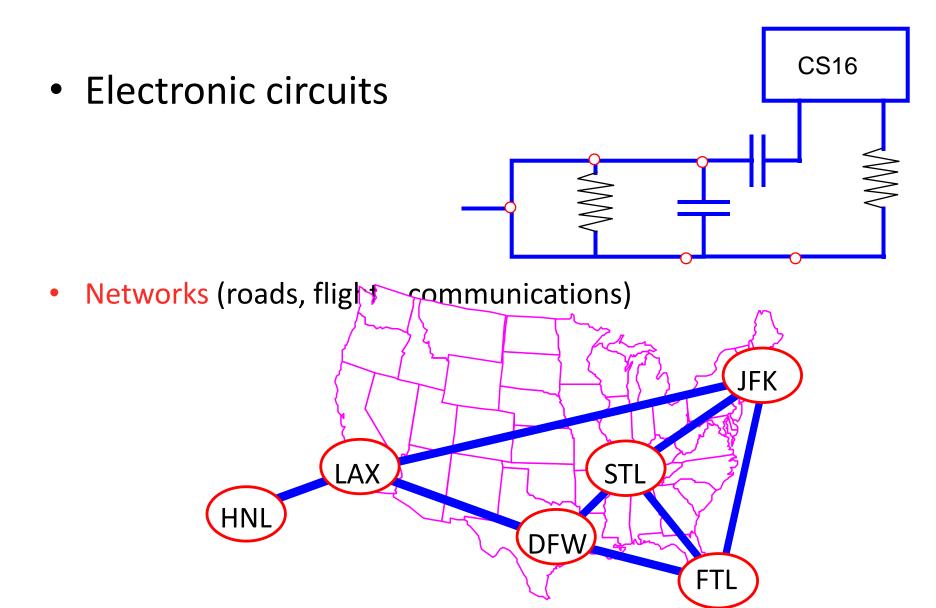
- An edge e = (u,v) is a pair of vertices
- Example:



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

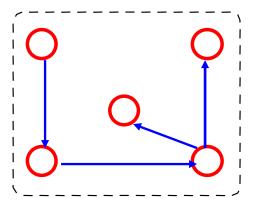
$$E= \{(a,b), (a,c), (a,d), (b,e), (c,d), (c,e), (d,e)\}$$

# **Applications**



# **Directed Graph**

A graph where edges are directed



## Directed vs. Undirected Graph

- An undirected graph is one in which the pair of vertices in an edge is unordered,  $(v_0, v_1) = (v_1, v_0)$
- A directed graph is one in which each edge is a directed pair of vertices,  $\langle v_0, v_1 \rangle \stackrel{!}{=} \langle v_1, v_0 \rangle$



# Terminology: Adjacent and Incident

- If  $(v_0, v_1)$  is an edge in an undirected graph,
  - −v₀ and v₁ are adjacent
  - The edge  $(v_0, v_1)$  is incident on vertices  $v_0$  and  $v_1$
- If (v<sub>0</sub>, v<sub>1</sub>) is an edge in a directed graph
  - $-v_0$  is adjacent to  $v_1$ , and  $v_1$  is adjacent from  $v_0$
  - The edge  $(v_0, v_1)$  is incident on  $v_0$  and  $v_1$

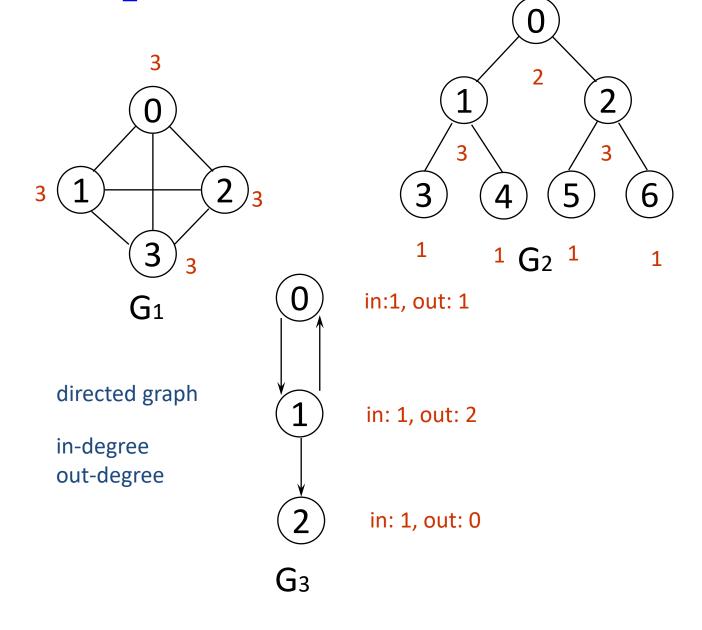
### Terminology: Degree of a Vertex

- The degree of a vertex is the number of edges incident to that vertex
- For directed graph,
  - the in-degree of a vertex v is the number of edges that have v as the head
  - the out-degree of a vertex v is the number of edges that have v as the tail
  - if  $d_i$  is the degree of a vertex i in a graph G with n vertices and e edges, the number of edges is

$$e = (\sum_{i=0}^{n-1} d_i) / 2$$

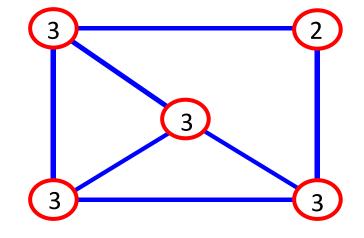
Why? Since adjacent vertices each count the adjoining edge, it will be counted twice

# **Examples**

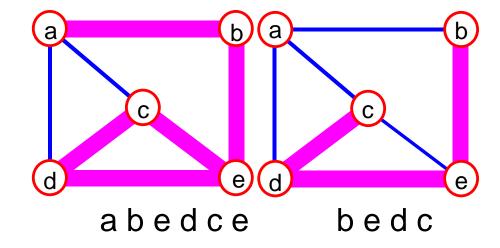


# Terminology: Path

path: sequence of vertices
 v<sub>1</sub>,v<sub>2</sub>,...v<sub>k</sub> such that
 consecutive vertices v<sub>i</sub> and
 v<sub>i+1</sub> are adjacent.

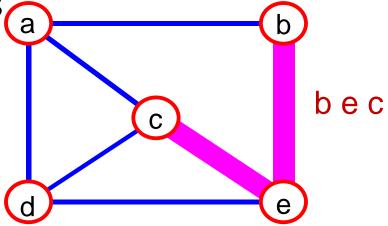


Not a PATH acbe



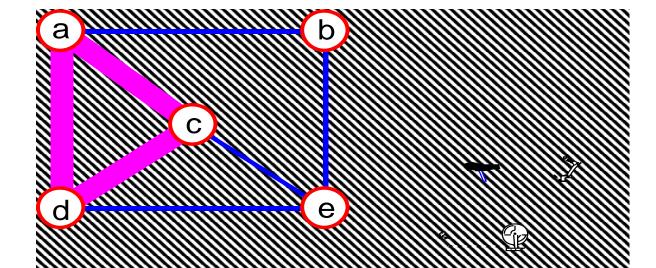
# More Terminology

• simple path: no repeated vertices



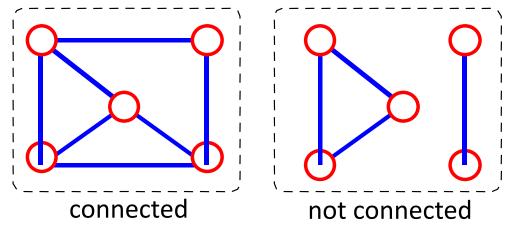
• cycle: simple path, except that the last vertex is the same as the first

vertex



# **Even More Terminology**

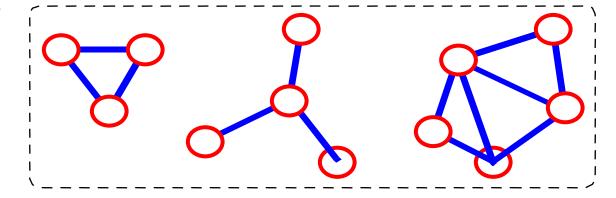
• Connected graph: any two vertices are connected by some path



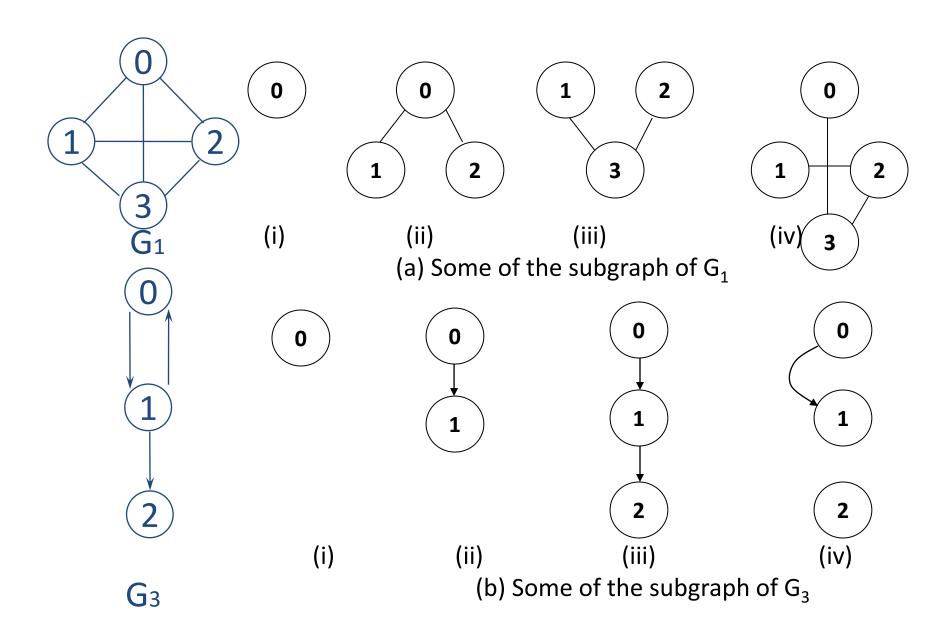
Subgraph: subset of vertices and edges forming a graph

Connected component: maximal connected subgraph. E.g., the graph below has 3

connected components.

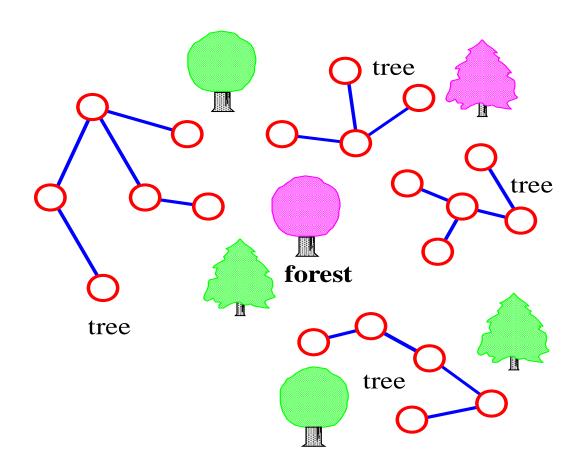


# Subgraphs Examples



#### More...

- tree connected graph without cycles
- forest collection of trees

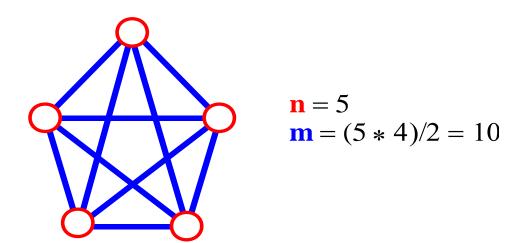


## Connectivity

- Let n = #vertices, and m = #edges
- A complete graph: one in which all pairs of vertices are adjacent
- How many total edges in a complete graph?

$$-m = n(n-1)/2.$$

• Therefore, if a graph is not complete, m < n(n-1)/2

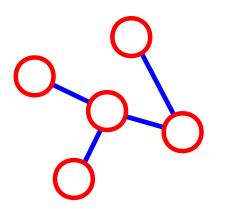


# More connectivity

n = #vertices

**m** = #edges

• For a tree **m** = **n** - 1



$$\mathbf{n} = 5$$

$$\mathbf{m} = 4$$

If  $\mathbf{m} < \mathbf{n} - 1$ , G is not connected

$$\begin{array}{c}
\mathbf{n} = 5 \\
\mathbf{m} = 3
\end{array}$$

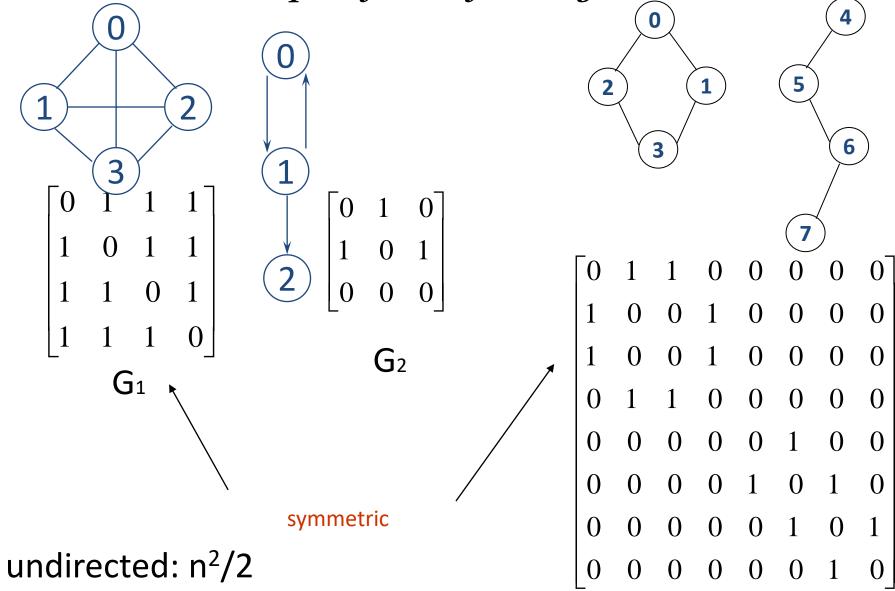
# **Graph Representations**

- Adjacency Matrix
- Adjacency Lists

# Data Structures for Graphs An Adjacency Matrix

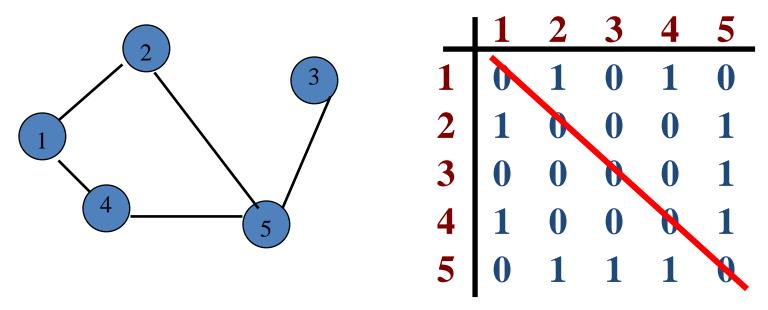
- Let G=(V,E) be a graph with n vertices.
- The adjacency matrix of G is a two-dimensional n by n array, say adj\_mat
- If the edge (v<sub>i</sub>, v<sub>j</sub>) is in E(G), adj\_mat[i][j]=1
- If there is no such edge in E(G), adj\_mat[i][j]=0

#### Examples for Adjacency Matrix



directed: n<sup>2</sup>

# **Adjacency Matrix Properties**



- Diagonal entries are zero
- The adjacency matrix of an undirected graph is symmetric;
   the adjacency matrix for a digraph need not be symmetric

# Adjacency Matrix

- The degree of a vertex i is  $\sum_{j=0}^{n-1} A[i][j]$
- For a digraph (= directed graph), the row sum is the out\_degree, while the column sum is the in\_degree of a vertex i

$$ind(v_i) = \sum_{j=0}^{n-1} A[j][i]$$
  $outd(v_i) = \sum_{j=0}^{n-1} A[i][j]$ 

## **Adjacency Matrix**

- n<sup>2</sup> bits of space
- All algos will require at least O(n²) time to find edges in G as n²-n
  entries of the matrix have to be examined (diagonal entries are zero)
- For an undirected graph, may store only lower or upper triangle (exclude diagonal)
  - -(n-1)n/2 bits
- O(n) time to find vertex degree and/or vertices adjacent to a given vertex
- Sparse graphs: problem
  - Speed up is possible through the use of linked lists in which only the edges that are in G are represented

# Data Structures for Graphs An Adjacency List

- A list of pointers, one for each node of the graph
- These pointers are the start of a linked list of nodes that can be reached by one edge of the graph
- For a weighted graph, this list would also include the weight for each edge

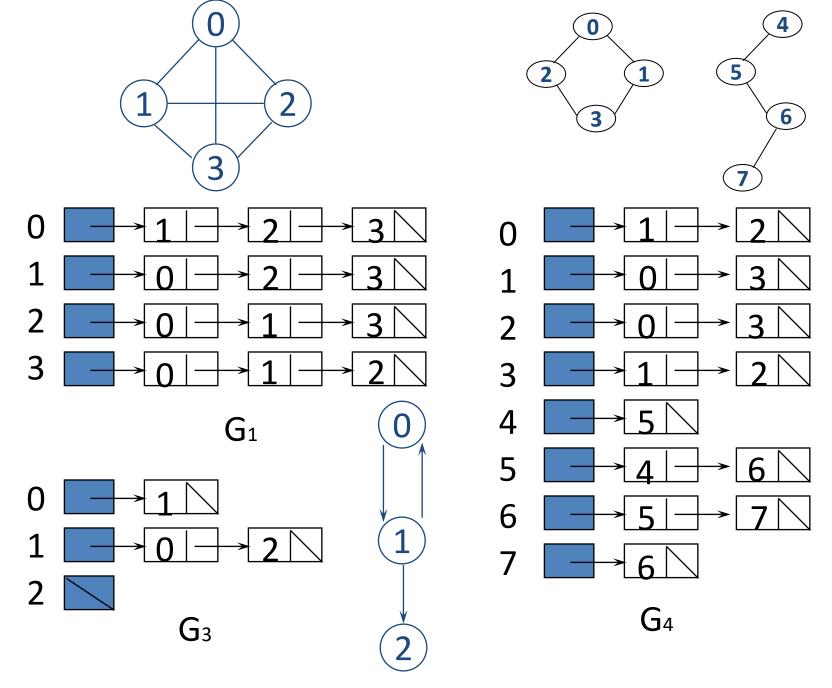
# Adjacency Lists (data structures)

Each row in adjacency matrix is represented as an adjacency list.

```
#define MAX_VERTICES 50

typedef struct node {
    int vertex_id;
        Node Structure
        struct node *link;
};

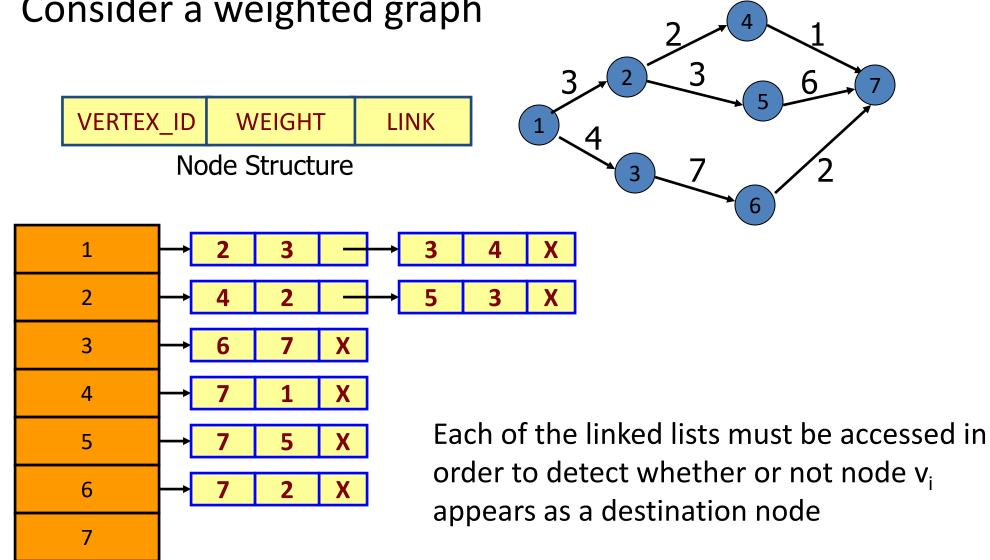
typedef struct node *node_pointer;
node_pointer graph[MAX_VERTICES];
```



An undirected graph with n vertices and e edges ==> n head nodes and 2e list nodes

# Adjacency Lists

Consider a weighted graph



# Adjacency Lists (data structures)

Each row in adjacency matrix is represented as an adjacency list.

```
#define MAX_VERTICES 50

typedef struct node {
    int vertex_id;
        Node Structure
        struct node *link;
};

typedef struct node *node_pointer;
node_pointer graph[MAX_VERTICES];
```

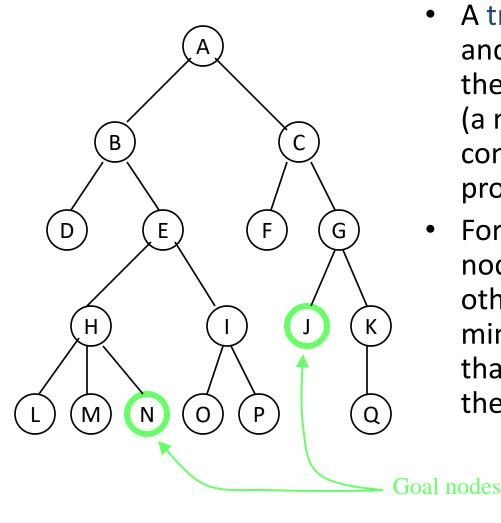
# **Some Operations**

- degree of a vertex in an undirected graph
  - # of nodes in its adjacency list
- # of edges in a graph
  - determined in O(v+e)
- out-degree of a vertex in a directed graph
  - # of nodes in its adjacency list
- in-degree of a vertex in a directed graph
  - traverse the whole data structure

#### **Graph Traversals**

- We want to travel to every node in the graph.
- Traversals guarantee that we will get to each node exactly once.
- This can be used if we want to search for information held in the nodes or if we want to distribute information to each node.

#### Tree searches



- A tree search starts at the root and explores nodes from there, looking for a goal node (a node that satisfies certain conditions, depending on the problem)
- For some problems, any goal node is acceptable (N or J); for other problems, you want a minimum-depth goal node, that is, a goal node nearest the root (only J)

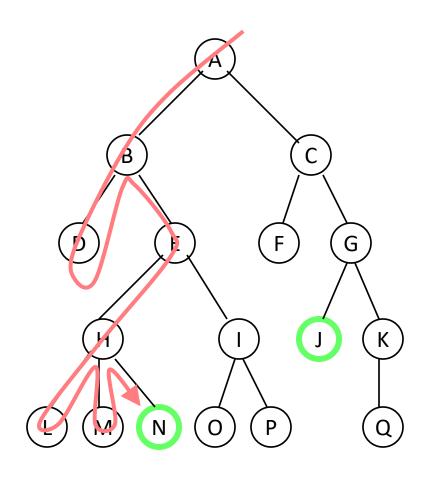
# **Graph Traversal**

- Problem: Search for a certain node or traverse all nodes in the graph
- Depth First Search (DFS)
  - Once a possible path is found, continue the search until the end of the path
- Breadth First Search (BFS)
  - Start several paths at a time, and advance in each one step at a time

#### **Depth-First Traversal**

- We follow a path through the graph until we reach a dead end.
- We then back up until we reach a node with an edge to an unvisited node
- We take this edge and again follow it until we reach a dead end
- This process continues until we back up to the starting node and it has no edges to unvisited nodes

# Depth-first searching in a Tree



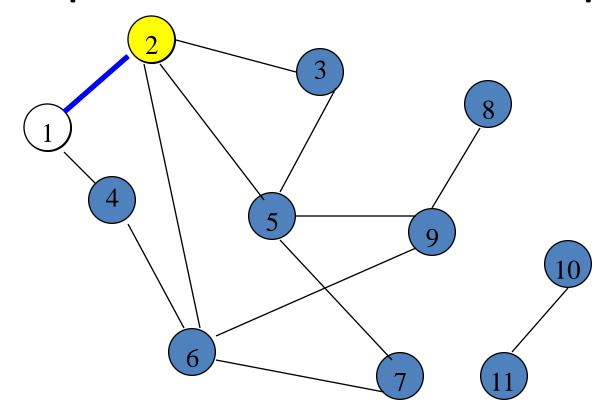
- A depth-first search (DFS)
   explores a path all the way to
   a leaf before backtracking and
   exploring another path
- For example, after searching A, then B, then D, the search backtracks and tries another path from B
- Node are explored in the order A B D E H L M N I O P C F G J K Q
- N will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

#### How to do DFS in a Tree

```
    Put the root node on a stack;
    while (stack is not empty) {
        pop a node from the stack;
        if (node is a goal node) return success;
            push all children of node onto the stack;
    }
    return failure;
```

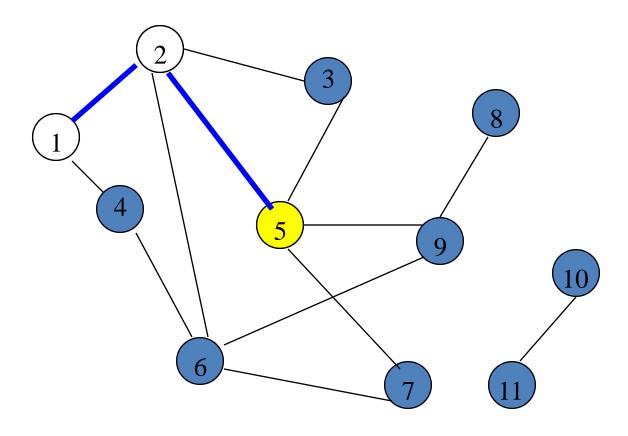
- At each step, the stack contains a path of nodes from the root
- The stack must be large enough to hold the longest possible path, that is, the maximum depth of search

## Depth-First Search Example



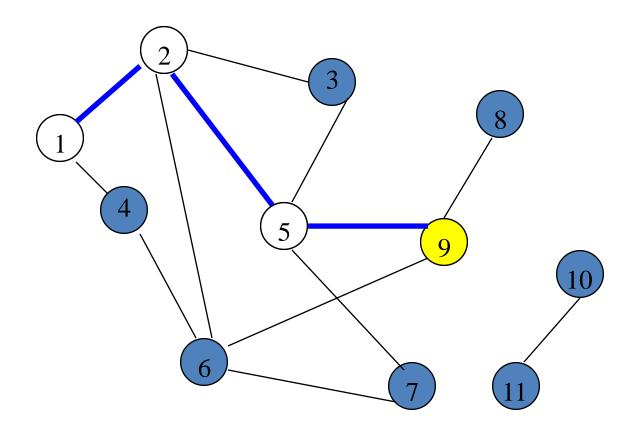
- Start search at vertex 1
- Label vertex 1 and do a depth first search from either 2 or 4
- Suppose that vertex 2 is selected

### Depth-First Search Example

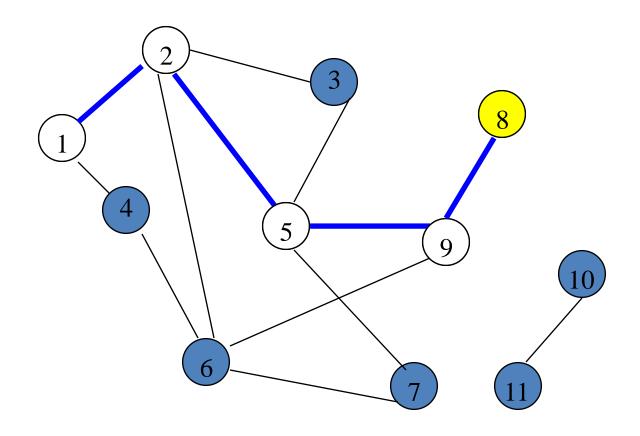


- Label vertex 2 and do a depth first search from either 3, 5, or 6
- Suppose that vertex 5 is selected

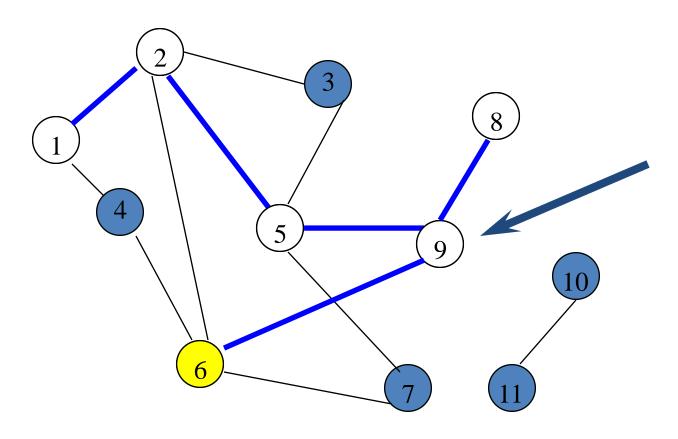
### Depth-First Search Example



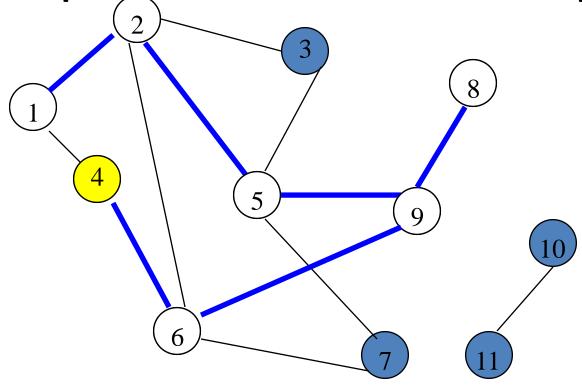
- Label vertex 5 and do a depth first search from either 3, 7, or 9
- Suppose that vertex 9 is selected



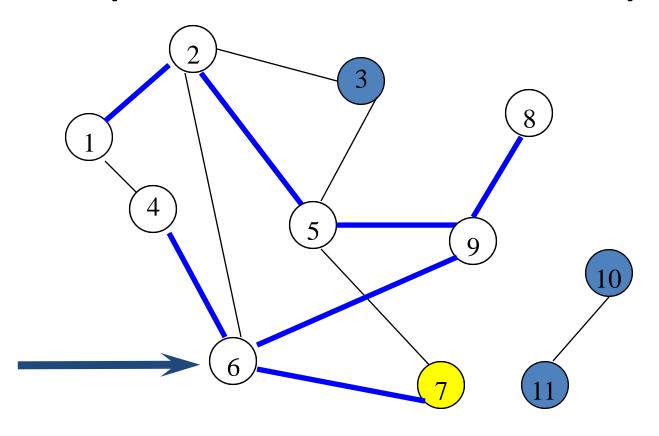
- Label vertex 9 and do a depth first search from either 6 or 8
- Suppose that vertex 8 is selected



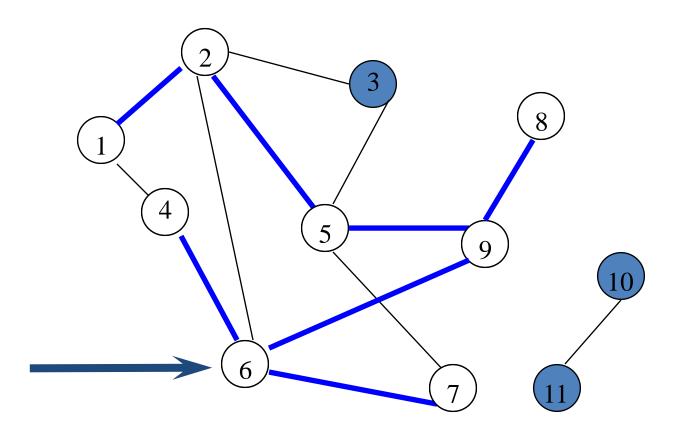
- Label vertex 8 and return to vertex 9
- From vertex 9 do a dfs(6)



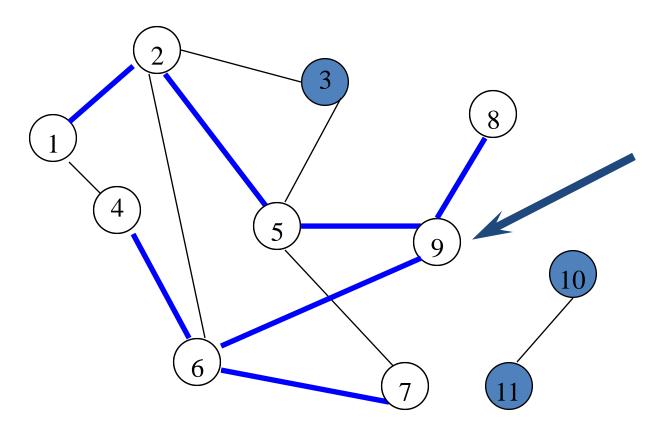
- Label vertex 6 and do a depth first search from either 4 or 7
- Suppose that vertex 4 is selected



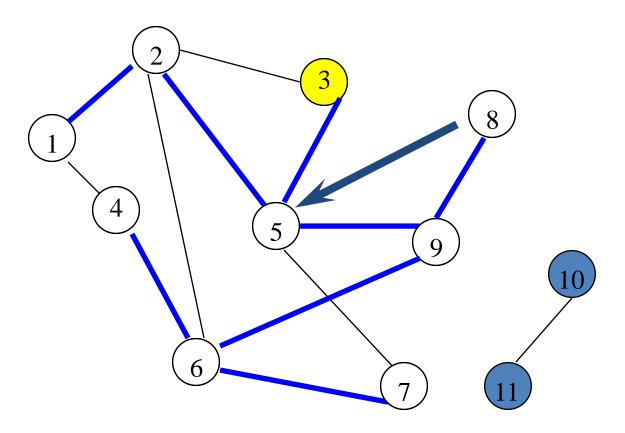
- Label vertex 4 and return to 6
- From vertex 6 do a dfs(7)



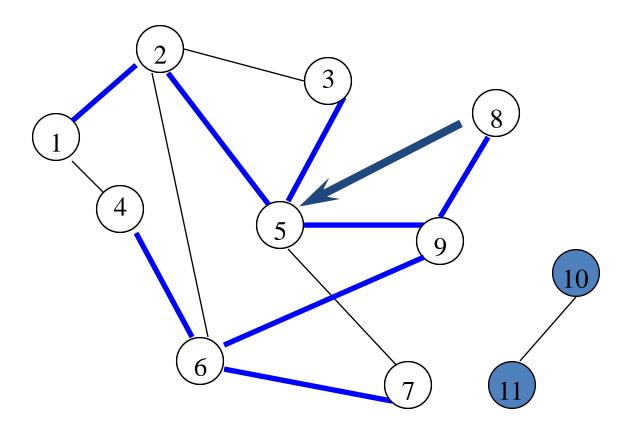
- Label vertex 7 and return to 6
- Return to 9



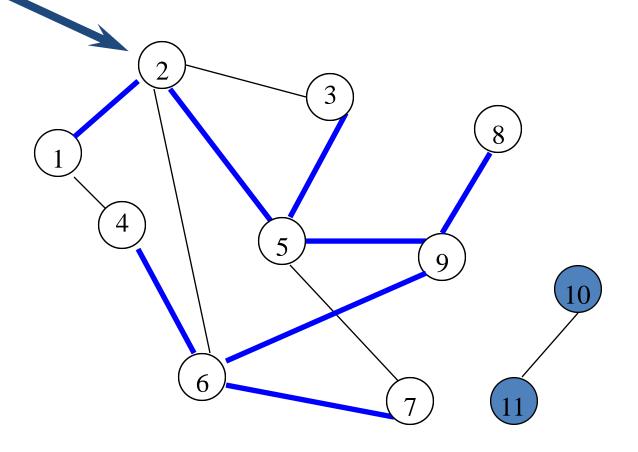
Return to 5



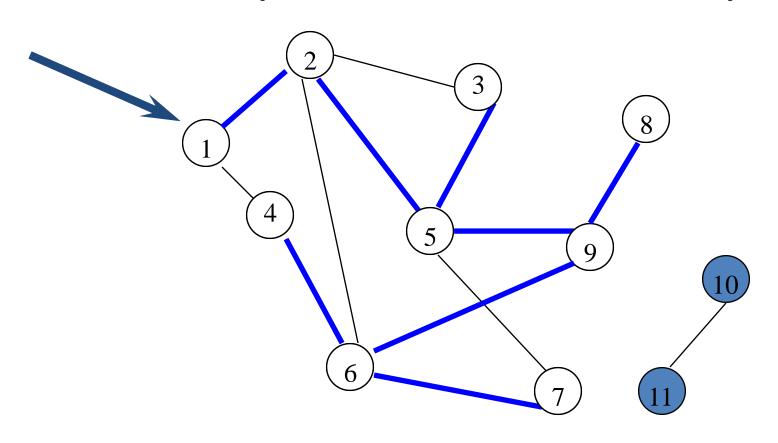
Do a dfs(3)



- Label 3 and return to 5
- Return to 2



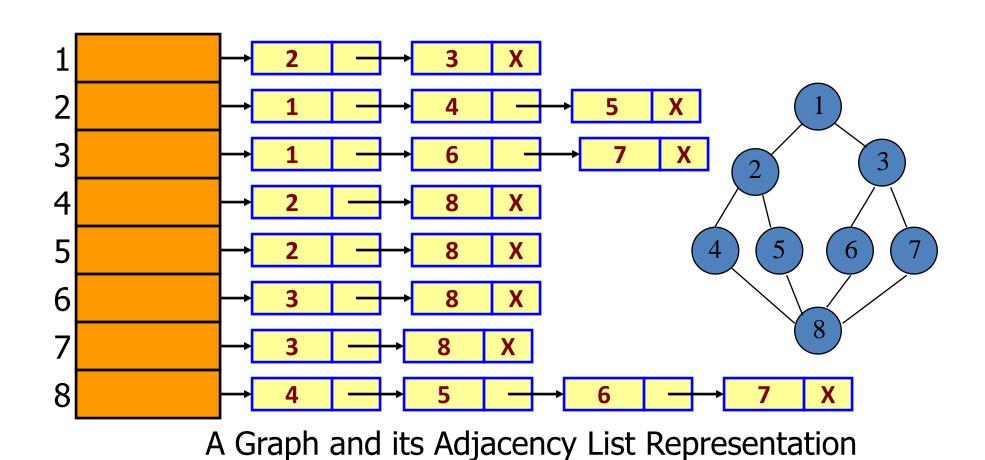
• Return to 1



Return to invoking method

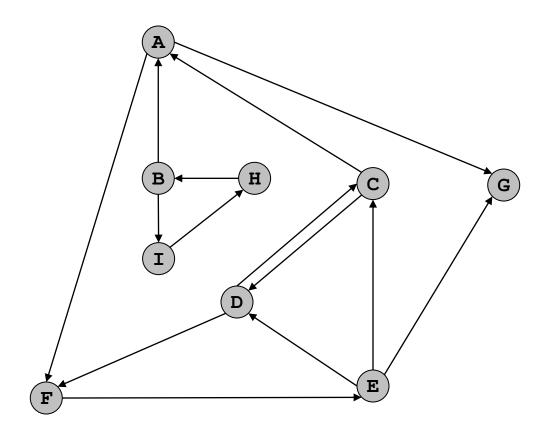
## Traversal: Another Example

• DFS (start vertex 1): 1, 2, 4, 8, 5, 6, 3, 7



### **DFS (Pseudo Code)**

```
DFS(input: Graph G) {
  Stack S; Integer x, t;
  while (G has an unvisited node x){
       visit(x); push(x,S);
       while (S is not empty){
               t := peek(S);
               if (t has an unvisited neighbor y){
                       visit(y); push(y,S); }
               else
                       pop(S);
```



#### Adjacency Lists

A: F 6

B: A I

C: A D

D: C F

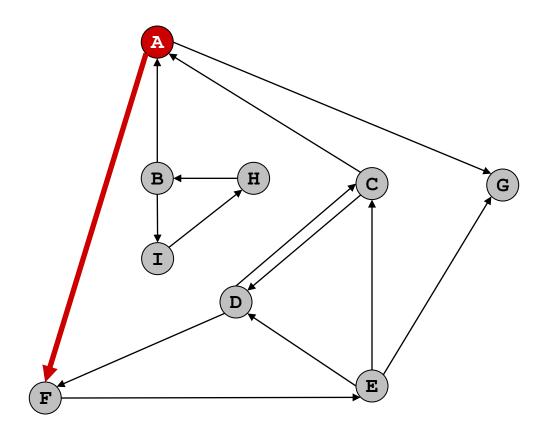
E: C D G

F: E

G:

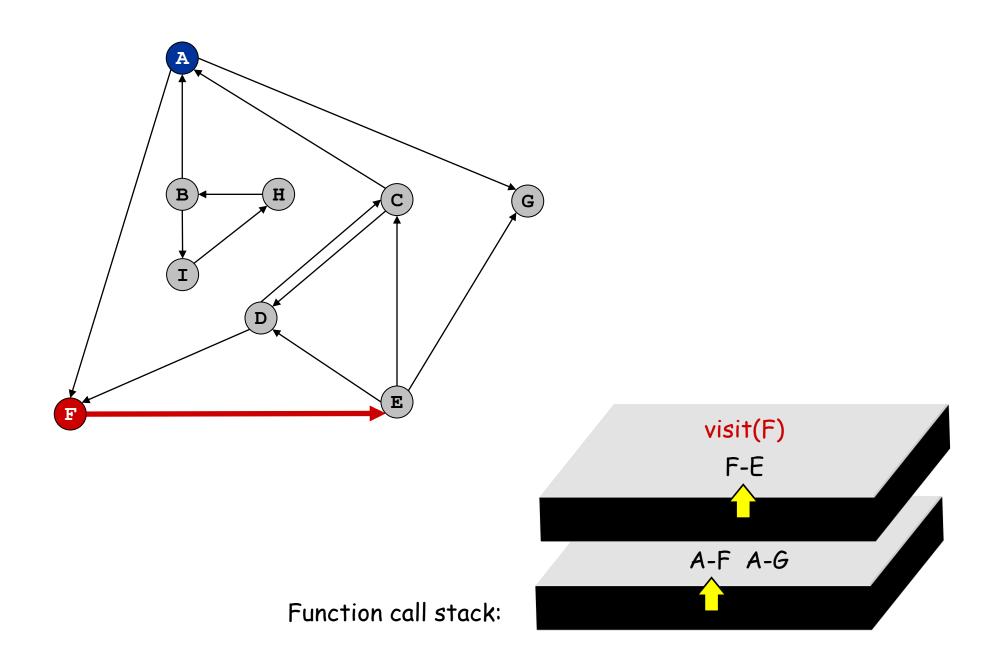
H: B

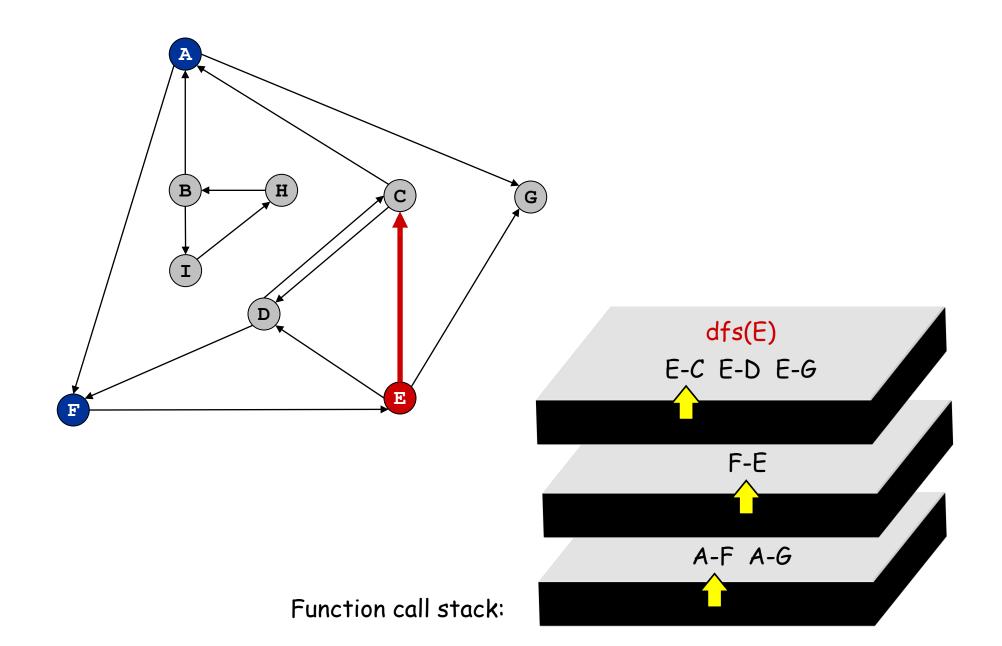
I: H

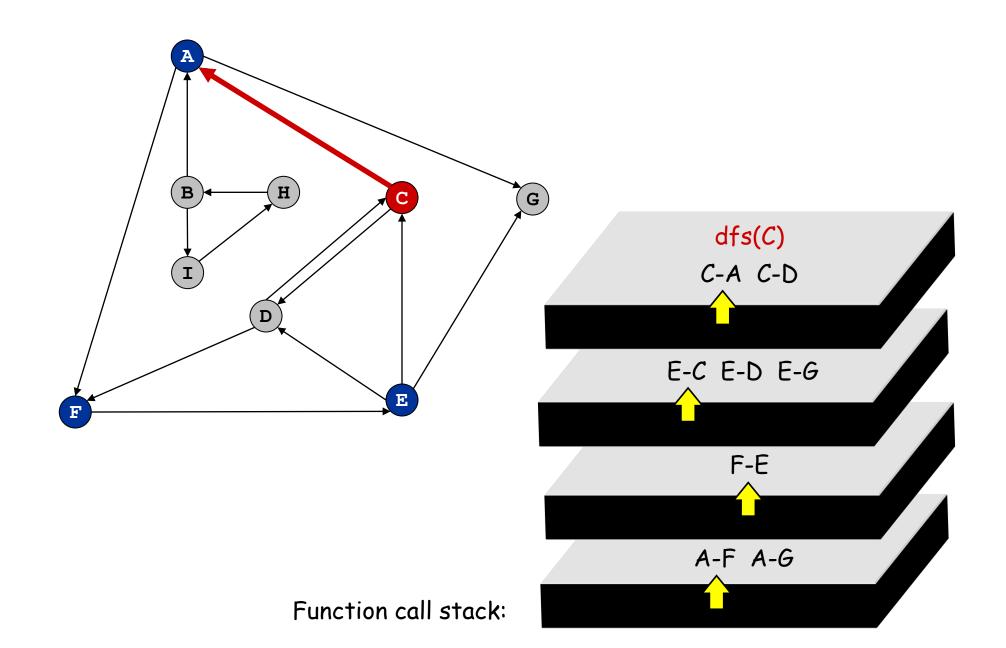


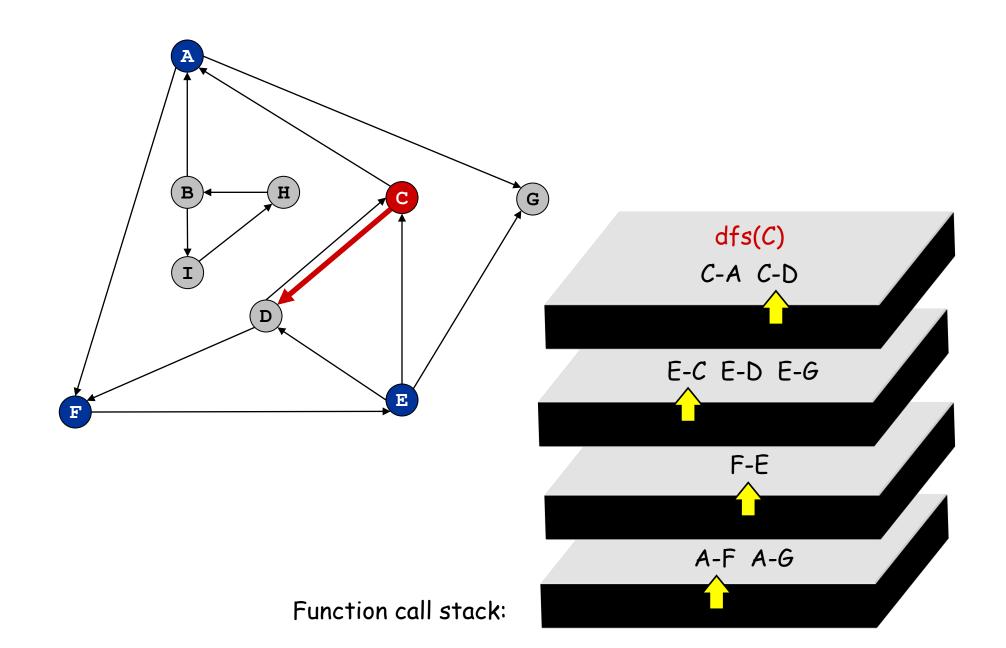


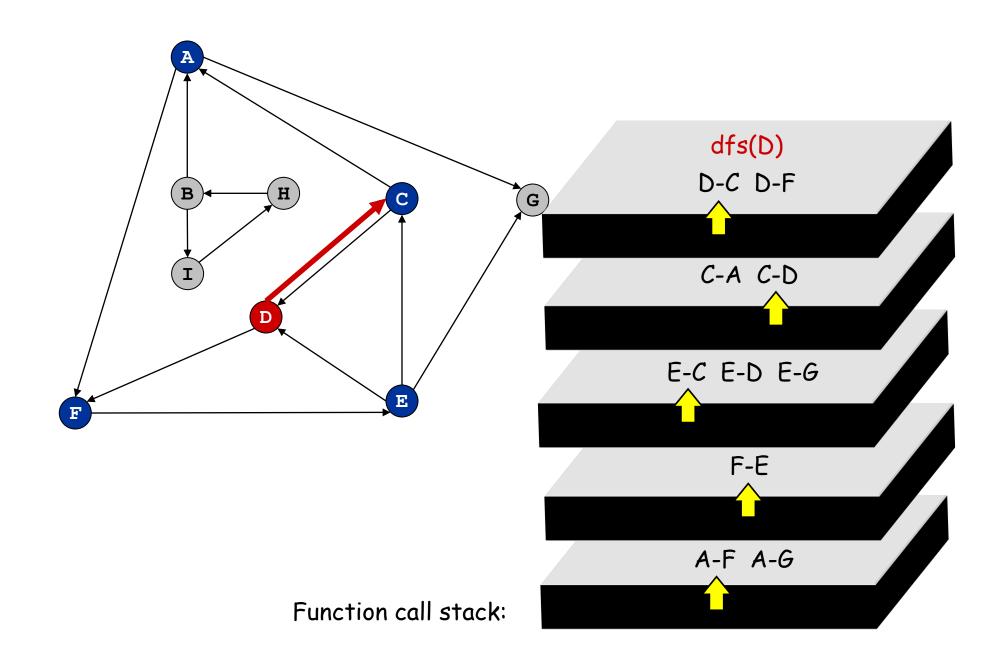
Function call stack:

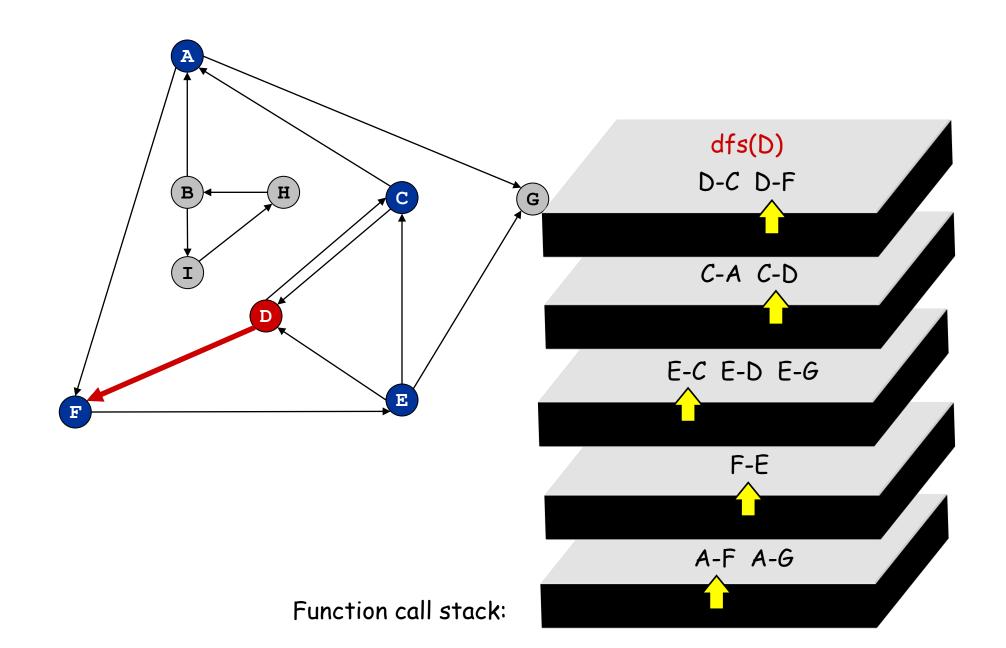


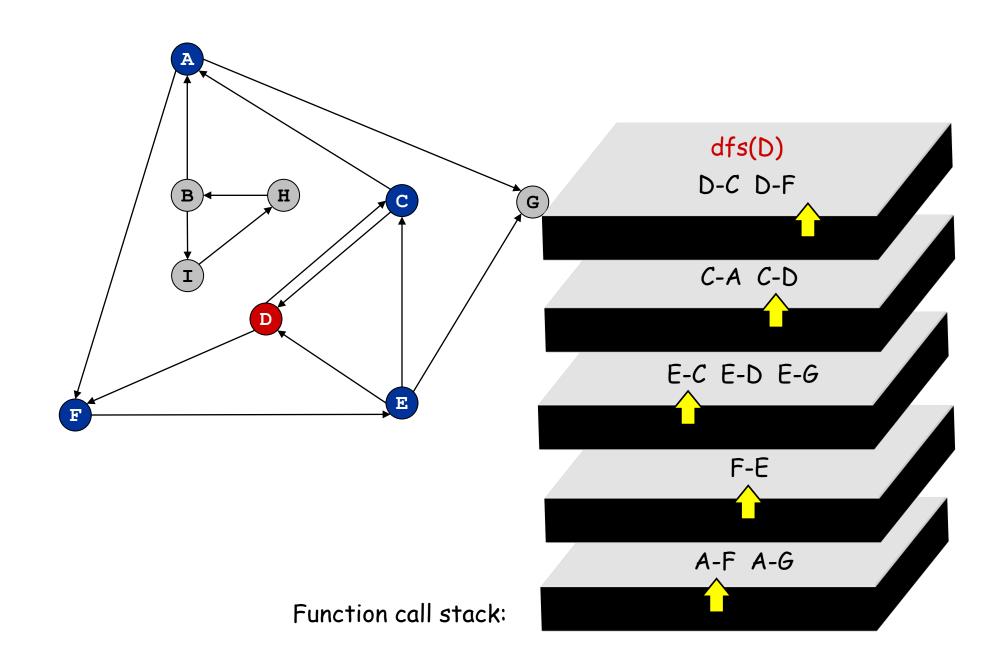


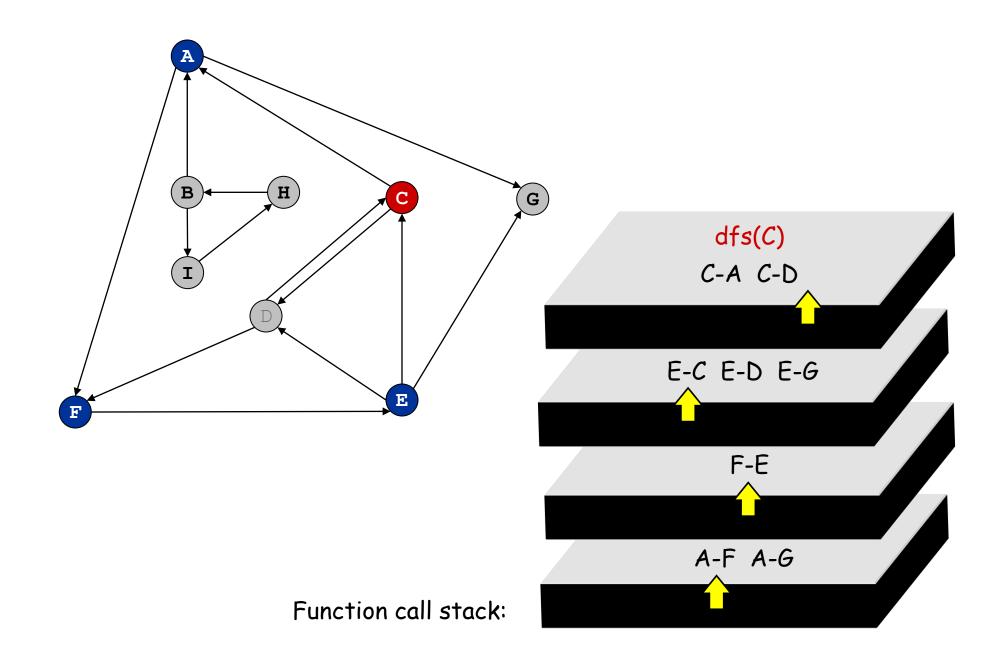


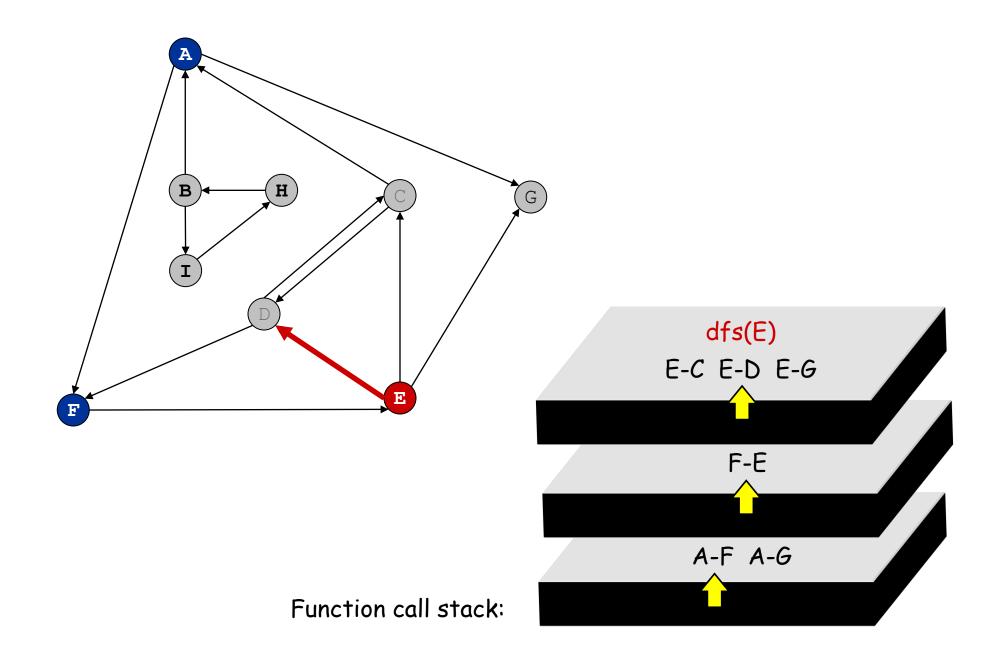


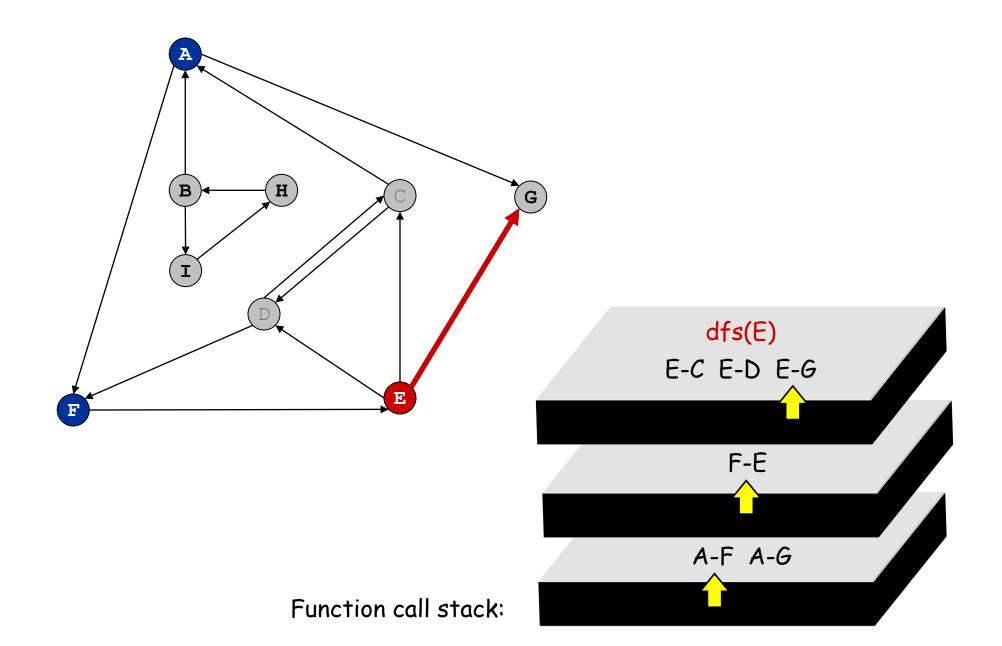


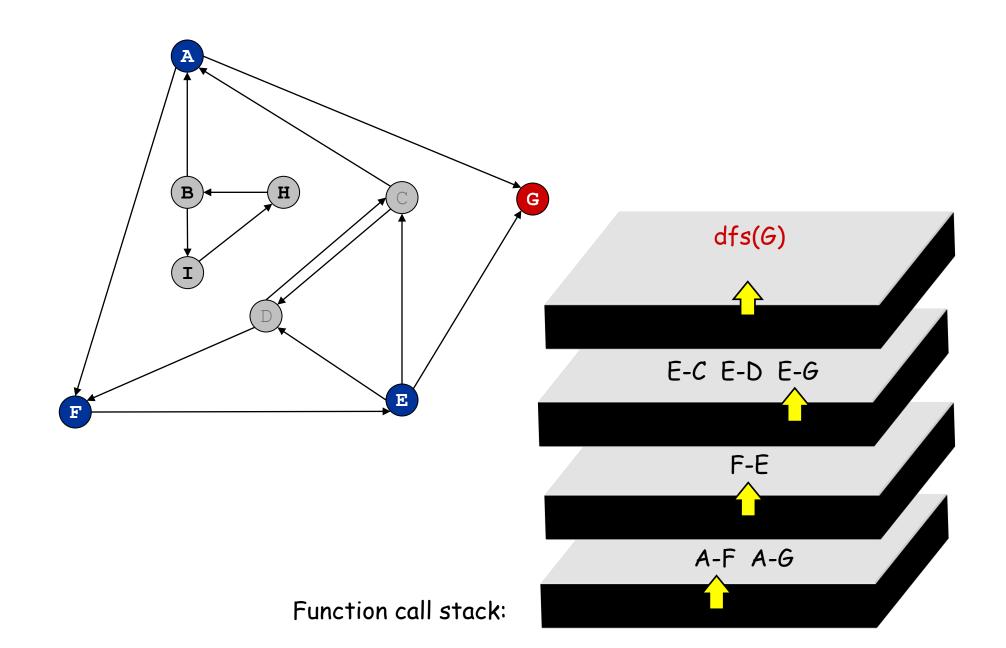


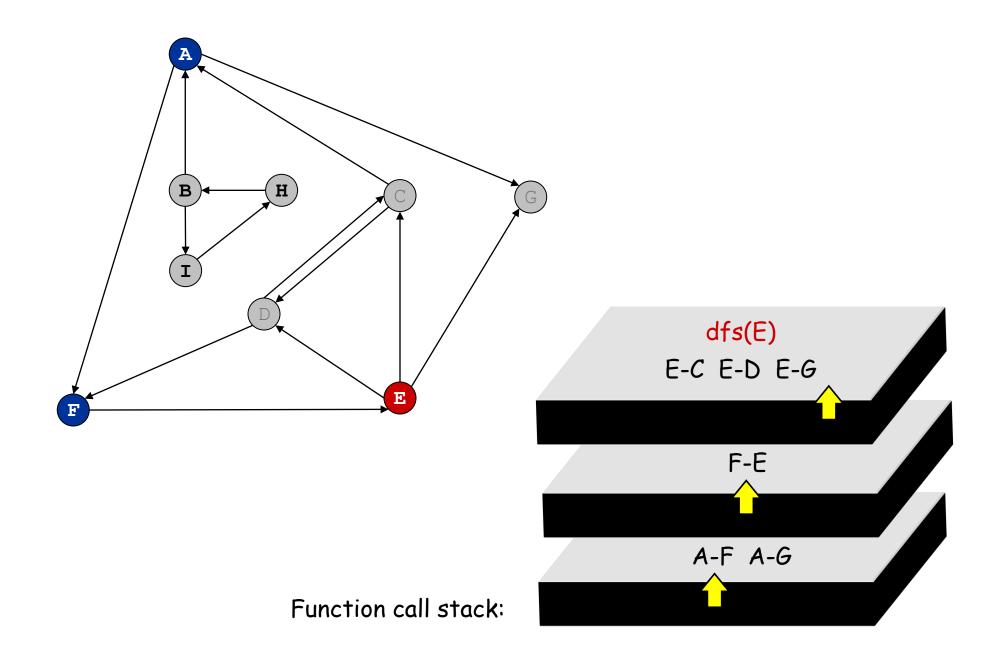


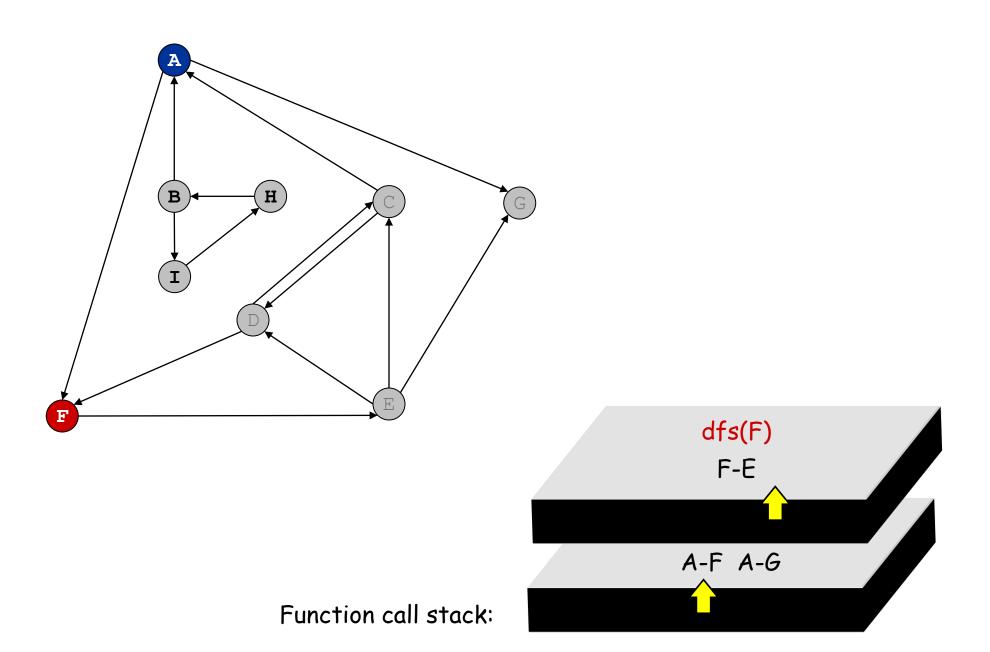


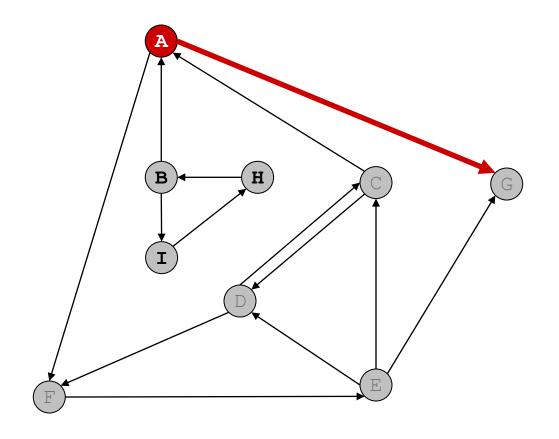


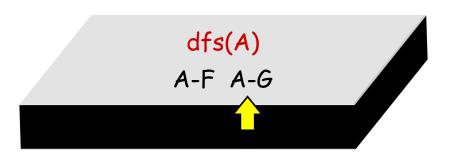




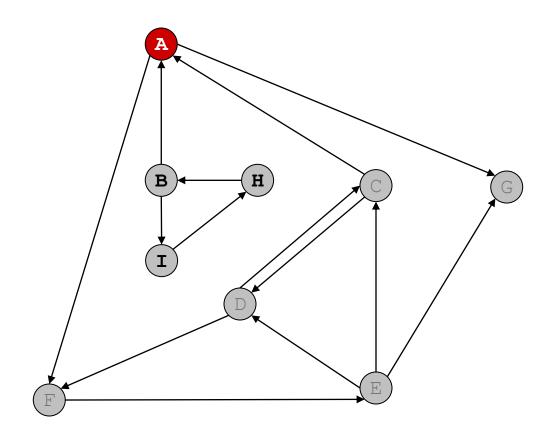






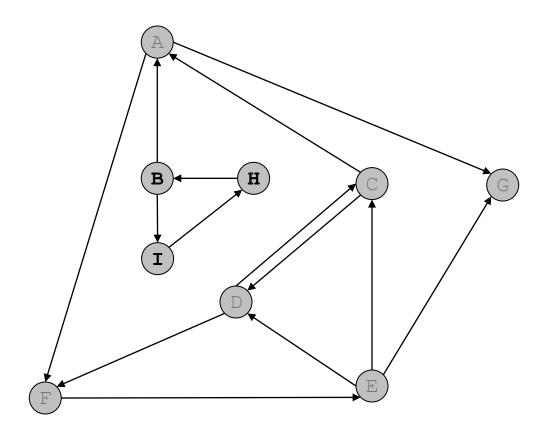


Function call stack:





Function call stack:



Nodes reachable from A: A, C, D, E, F, G

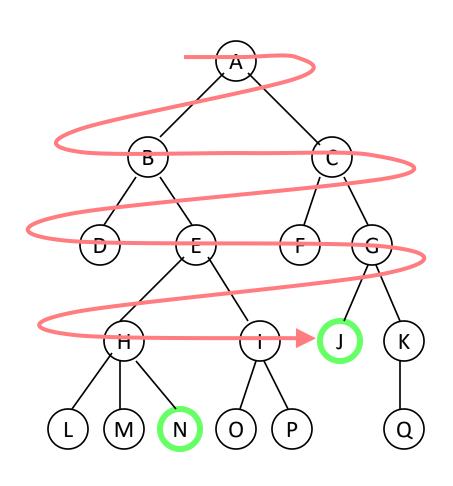
#### **Breadth-First Traversal**

- From the starting node, we follow all paths of length one
- Then we follow paths of length two that go to unvisited nodes
- We continue increasing the length of the paths until there are no unvisited nodes along any of the paths

### **Breadth-First Search**

- Visit start vertex and put into a FIFO queue
- Repeatedly remove a vertex from the queue, visit its unvisited adjacent vertices, put newly visited vertices into the queue

### Breadth-first searching in a Tree



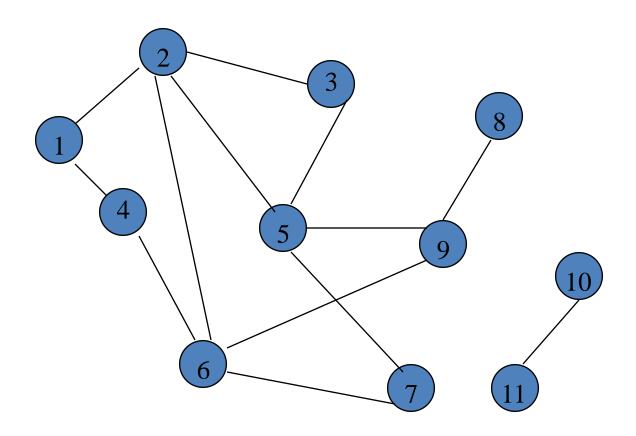
- A breadth-first search (BFS)
   explores nodes nearest the root
   before exploring nodes further
   away
- For example, after searching A, then B, then C, the search proceeds with D, E, F, G
- Node are explored in the order A
   B C D E F G H I J K L M N O P
   Q
- J will be found before N

### How to do BFS in a Tree

```
    Put the root node on a queue;
    while (queue is not empty) {
        remove a node from the queue;
        if (node is a goal node) return success;
            put all children of node onto the queue;
        }
        return failure;
```

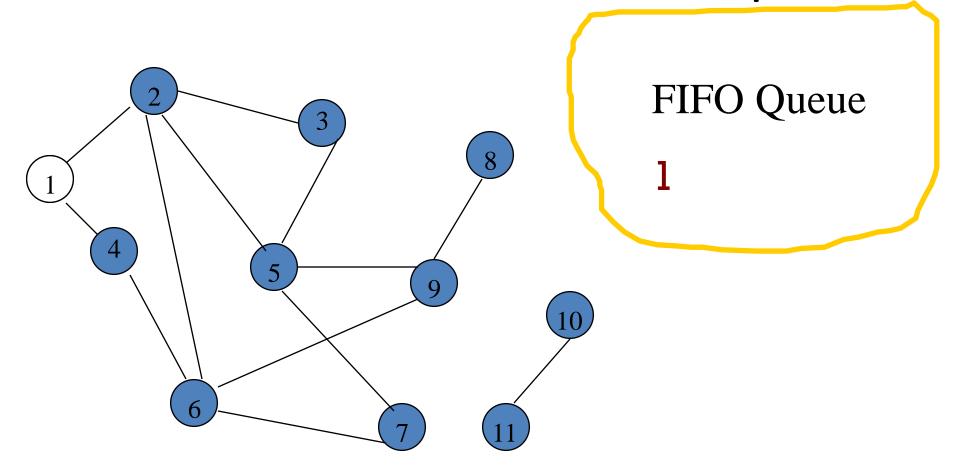
- Just before starting to explore level i, the queue holds all the nodes at level i-1
- In a typical tree, the number of nodes at each level increases exponentially with the depth
- Memory requirements may be infeasible
- When this method succeeds, it doesn't give the path

# Breadth-First Search Example

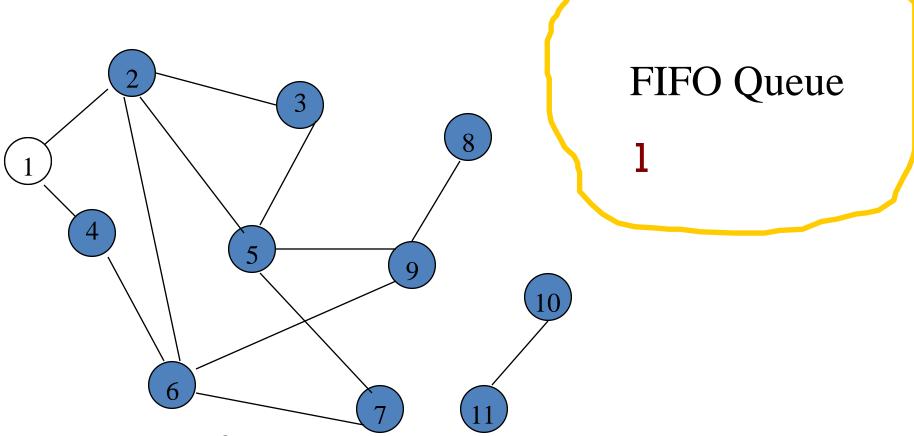


Start search at vertex 1

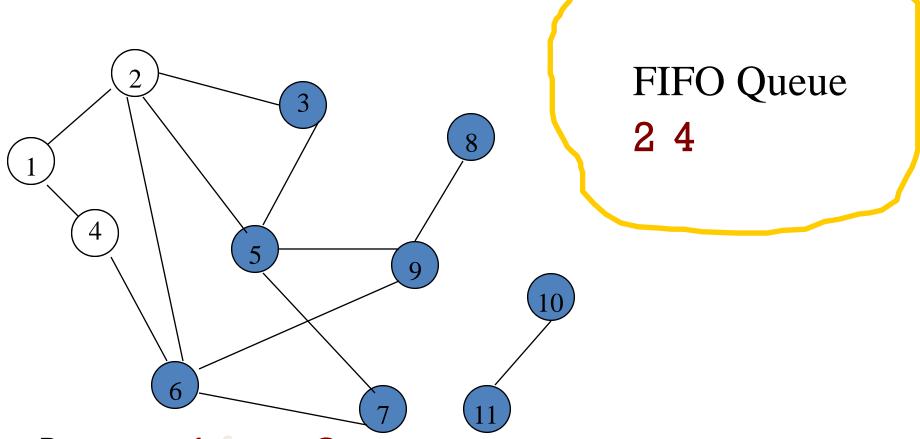
# Breadth-First Search Example



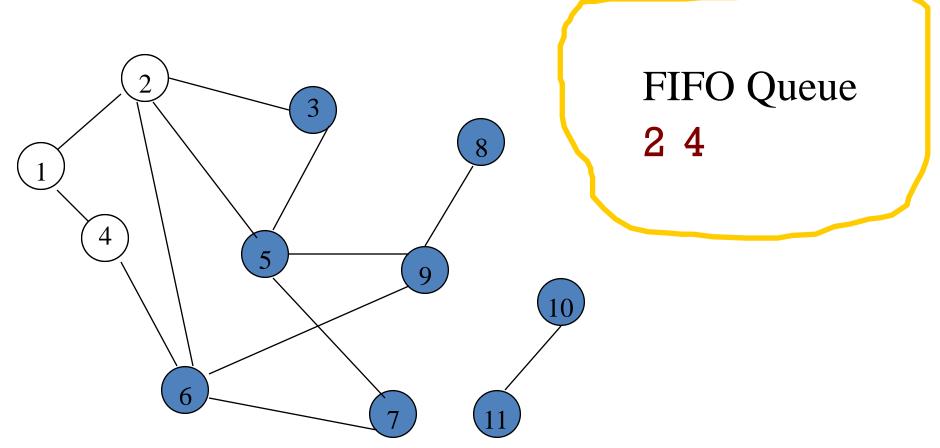
Visit/mark/label start vertex and put in a FIFO queue



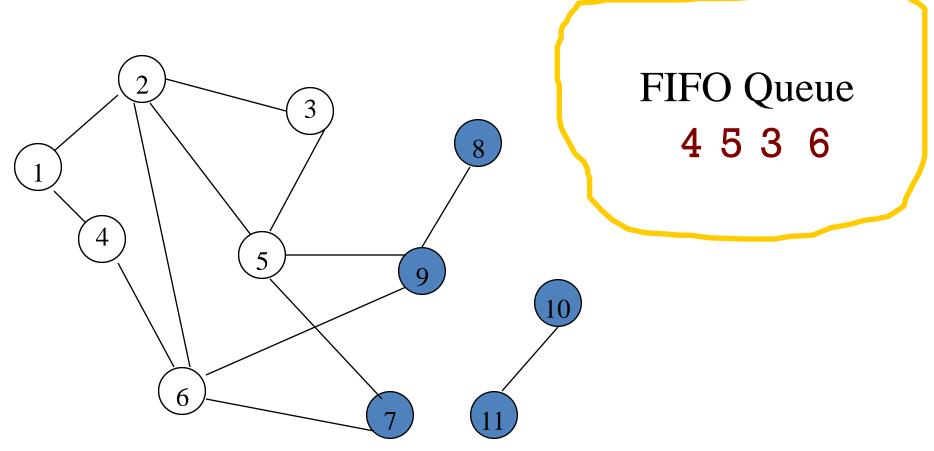
- Remove 1 from Q
- Visit adjacent unvisited vertices & put them in Q



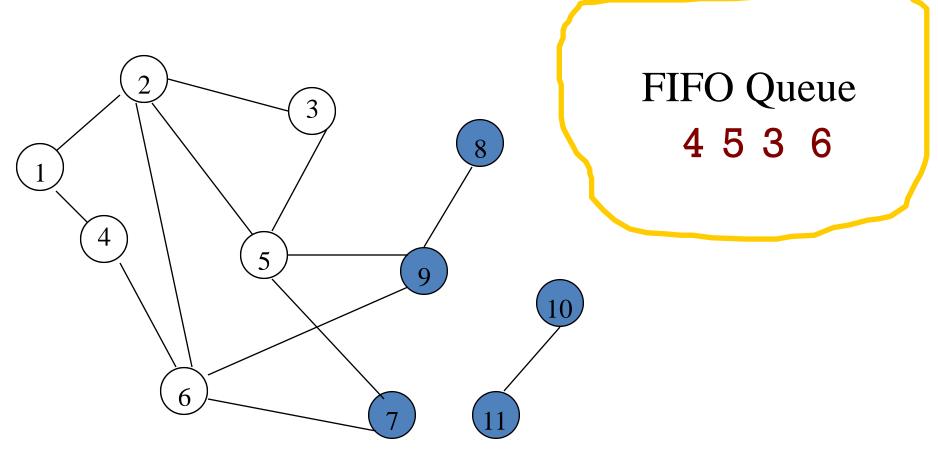
- Remove 1 from Q
- Visit adjacent unvisited vertices & put them in Q



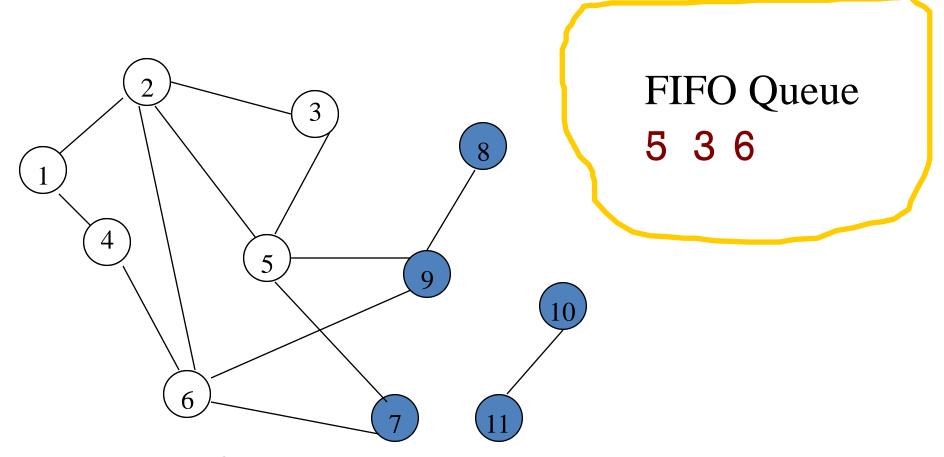
- Remove 2 from Q
- Visit adjacent unvisited vertices & put them in



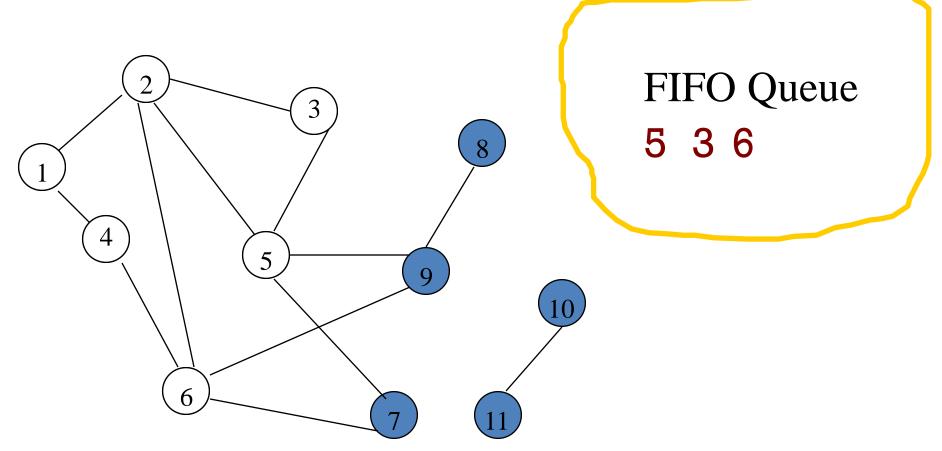
- Remove 2 from Q
- Visit adjacent unvisited vertices & put them in Q



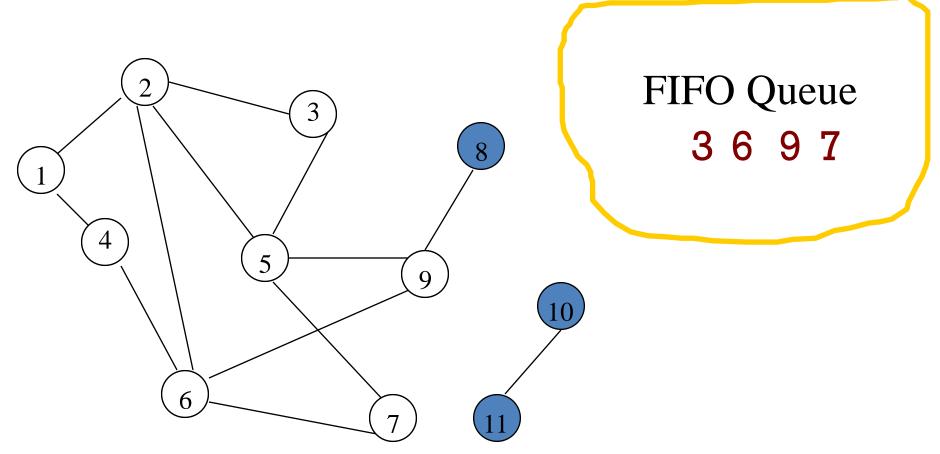
- Remove 4 from Q
- Visit adjacent unvisited vertices & put them in Q



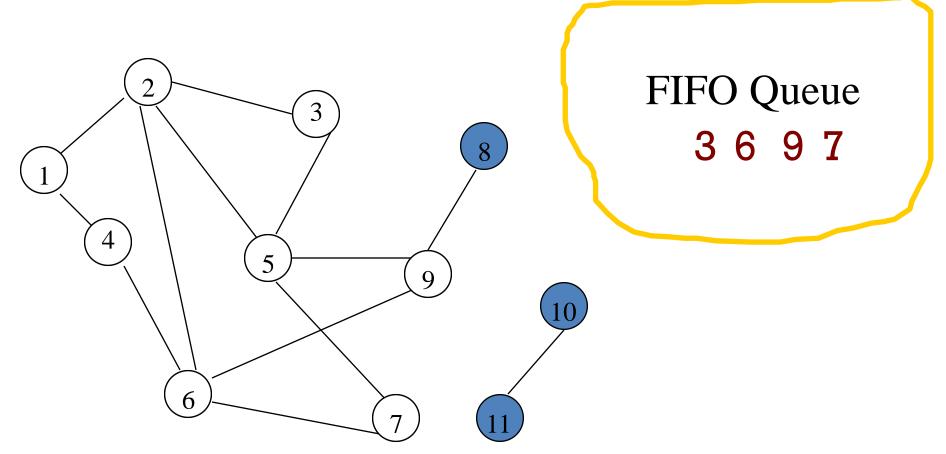
- Remove 4 from Q
- Visit adjacent unvisited vertices & put them in Q



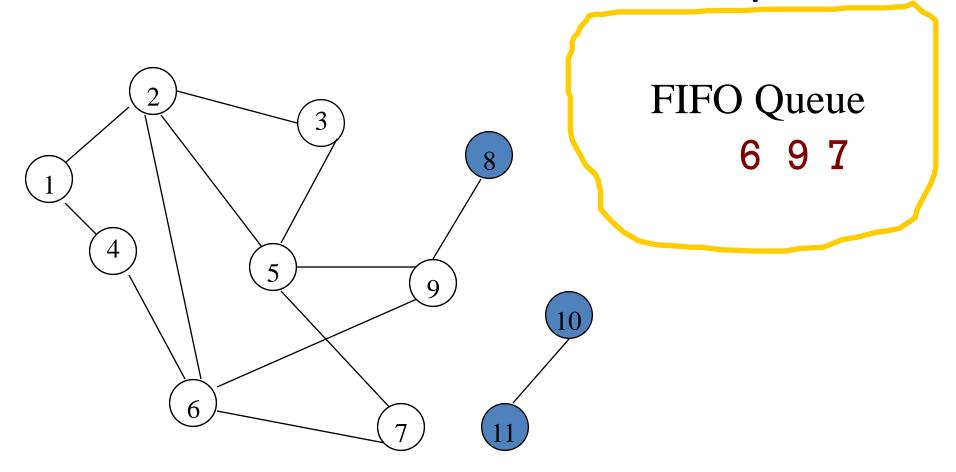
- Remove 5 from Q
- Visit adjacent unvisited vertices & put them in Q



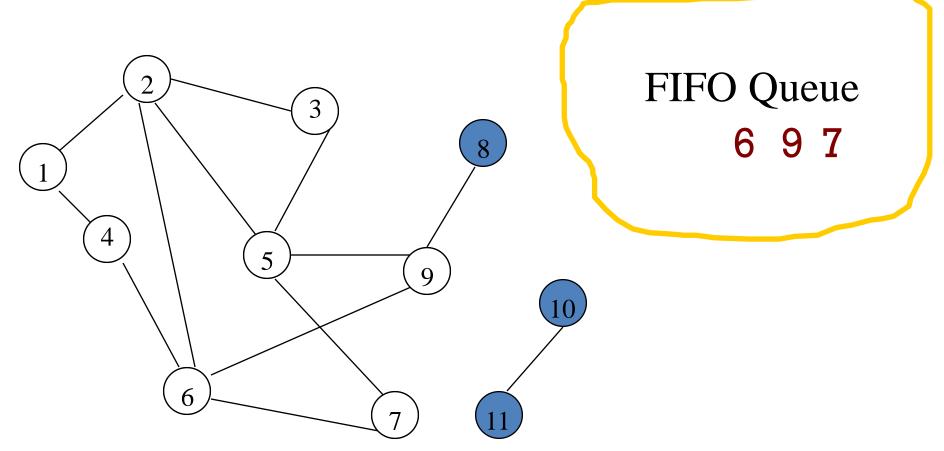
- Remove 5 from Q
- Visit adjacent unvisited vertices & put them in Q



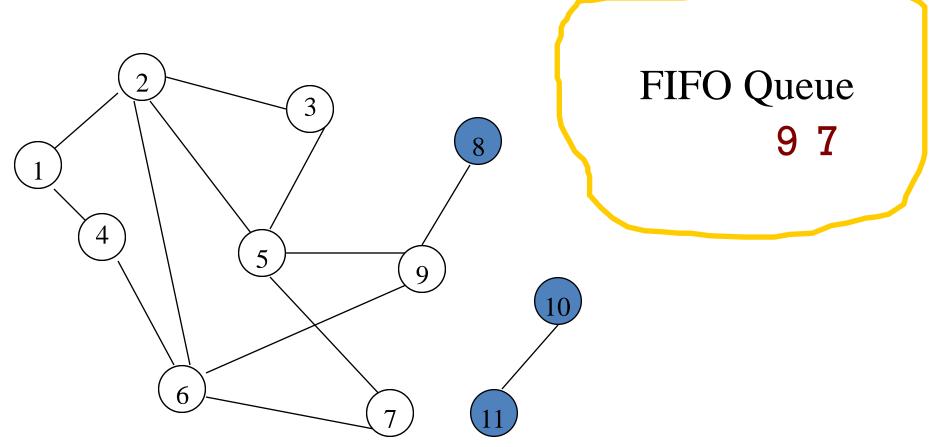
- Remove 3 from Q
- Visit adjacent unvisited vertices & put them in Q



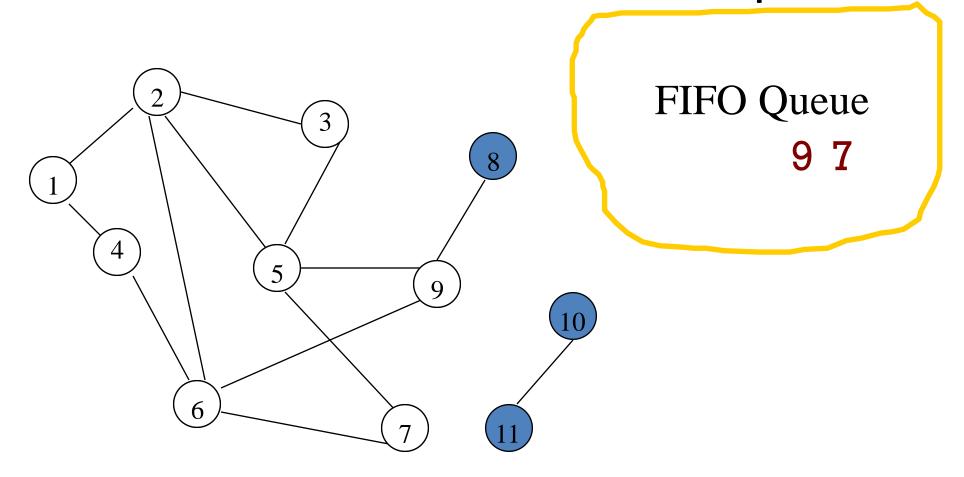
- Remove 3 from Q
- Visit adjacent unvisited vertices & put them in Q



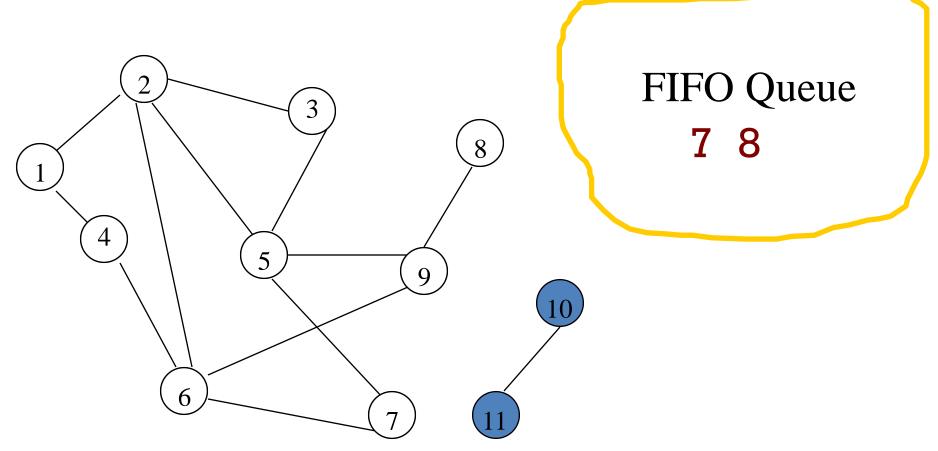
- Remove 6 from Q
- Visit adjacent unvisited vertices & put them in Q



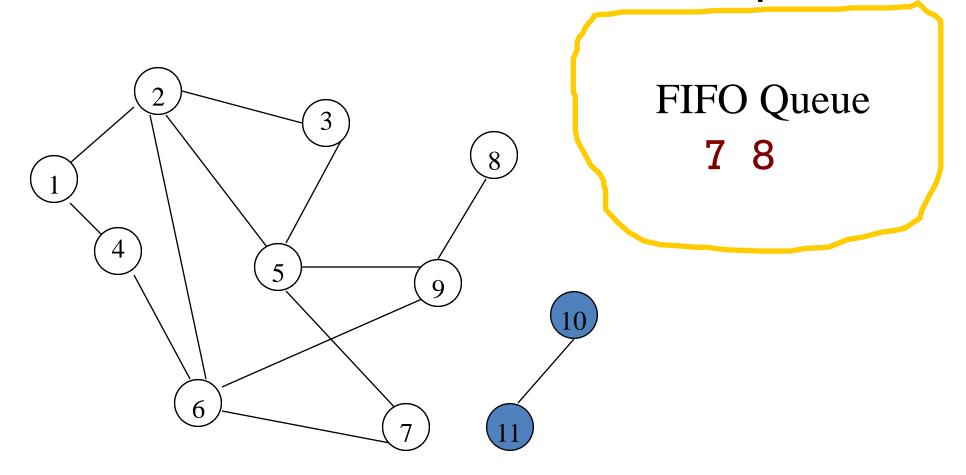
- Remove 6 from Q
- Visit adjacent unvisited vertices & put them in Q



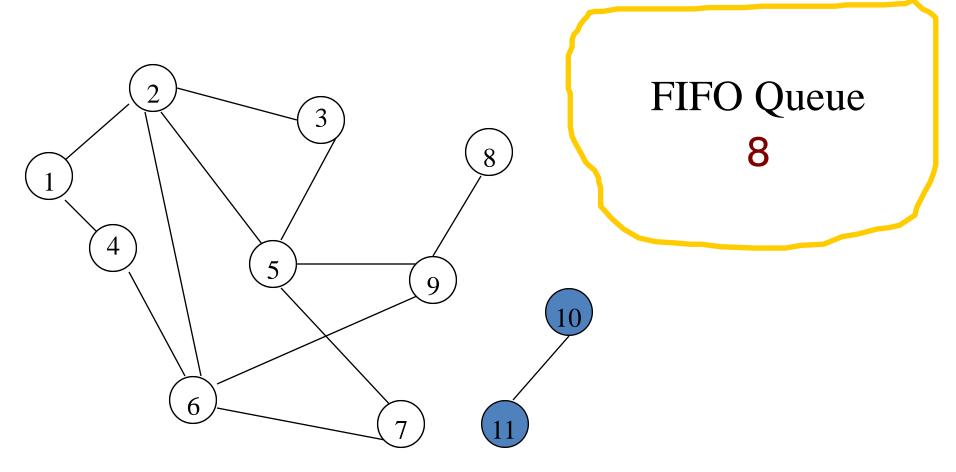
- Remove 9 from Q
- Visit adjacent unvisited vertices & put them in Q



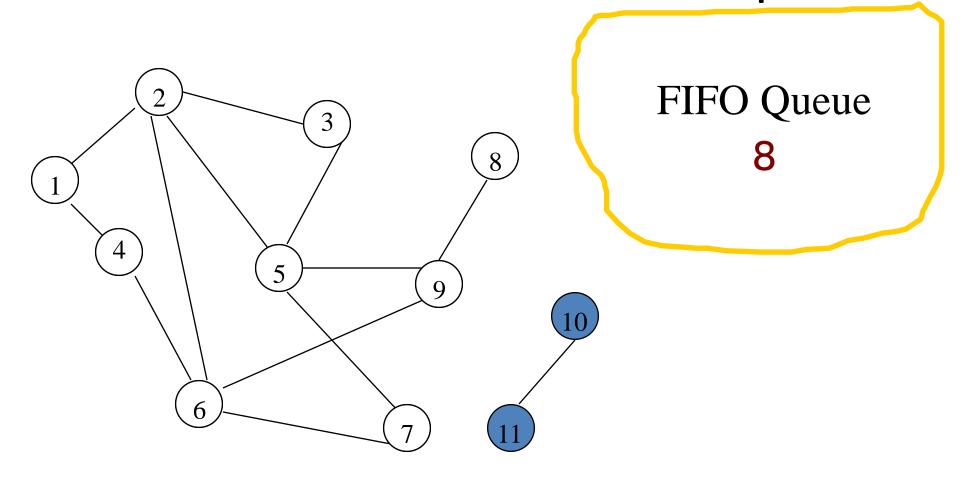
- Remove 9 from Q
- Visit adjacent unvisited vertices & put them in Q



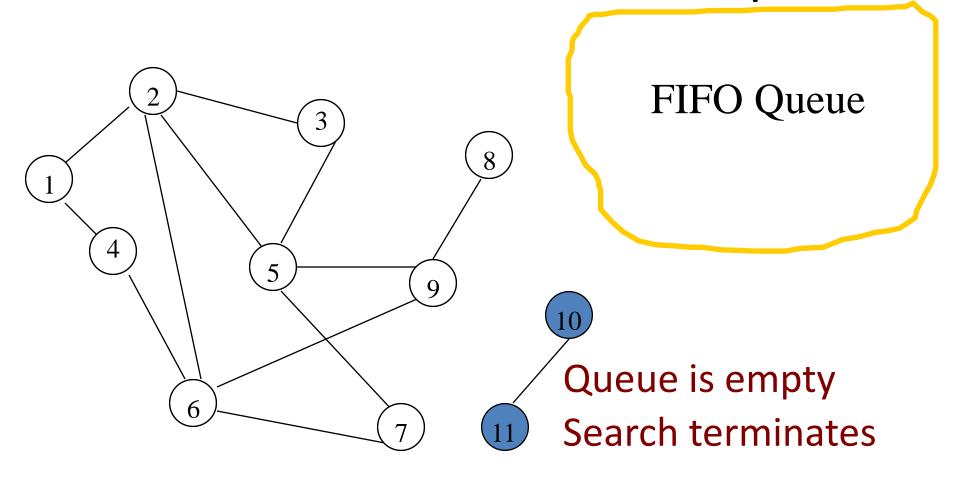
- Remove 7 from Q
- Visit adjacent unvisited vertices & put them in Q



- Remove 7 from Q
- Visit adjacent unvisited vertices & put them in Q

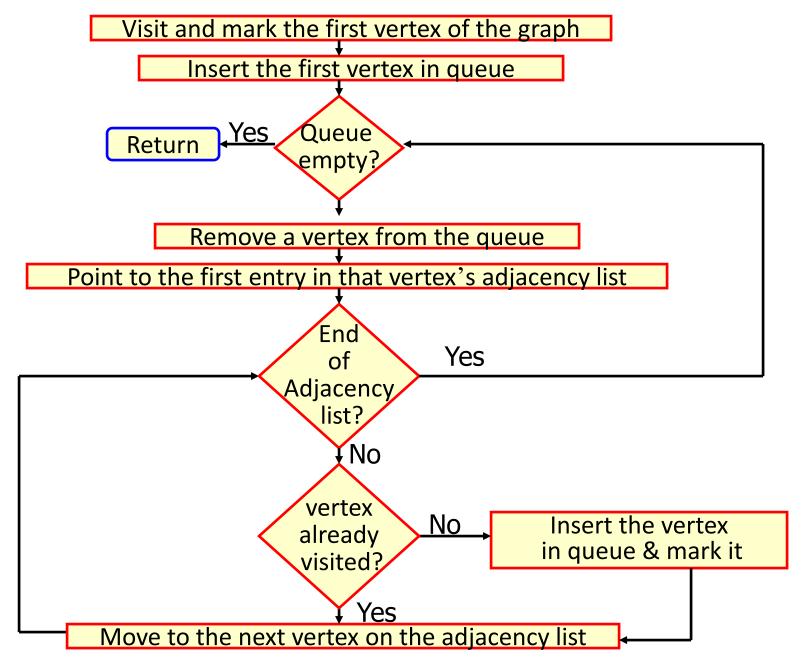


- Remove 8 from Q
- Visit adjacent unvisited vertices & put them in Q



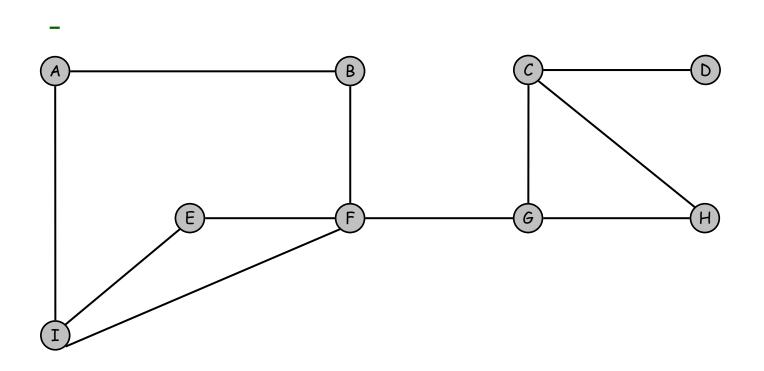
■ All vertices reachable from the start vertex (including the start vertex) are visited

### **BFS- Flowchart**



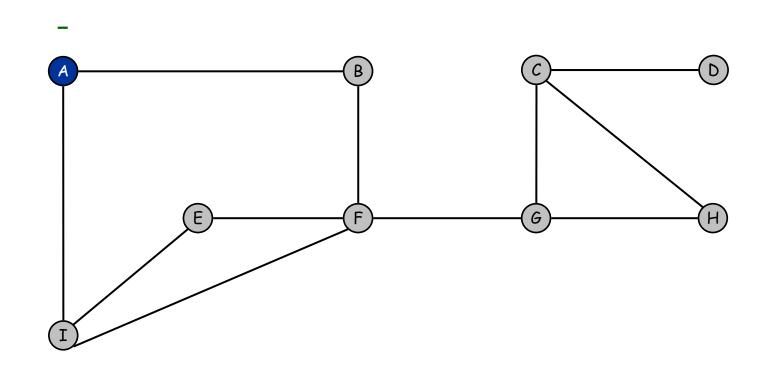
### **BFS (Pseudo Code)**

```
BFS(input: graph G) {
  Queue Q; Integer x, z, y;
  while (G has an unvisited node x) {
       visit(x); Enqueue(x,Q);
       while (Q is not empty){
               z := Dequeue(Q);
               for all (unvisited neighbor y of z){
                      visit(y); Enqueue(y,Q);
```



front

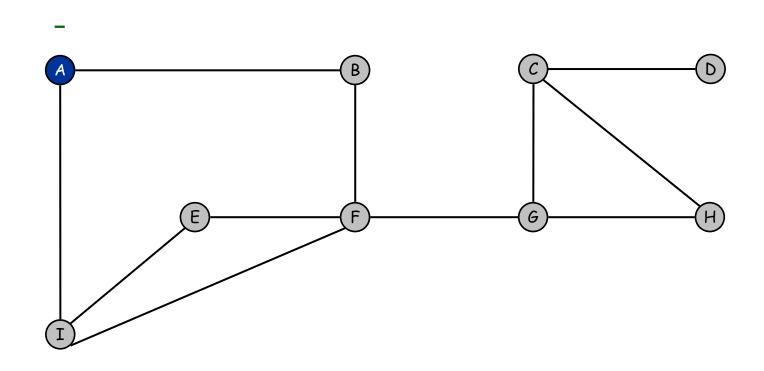
FIFO Queue



enqueue source node

front

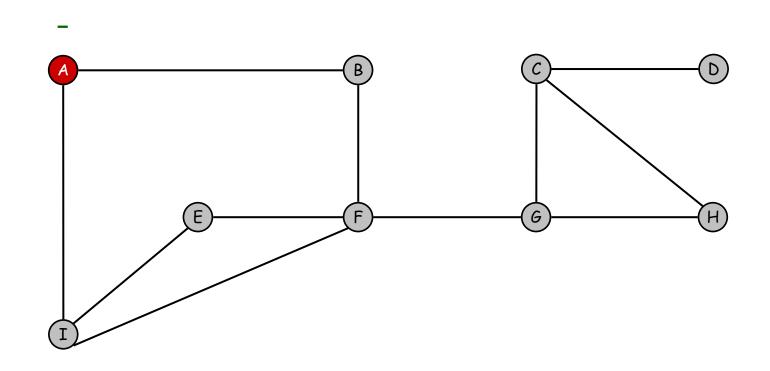
A



dequeue next vertex

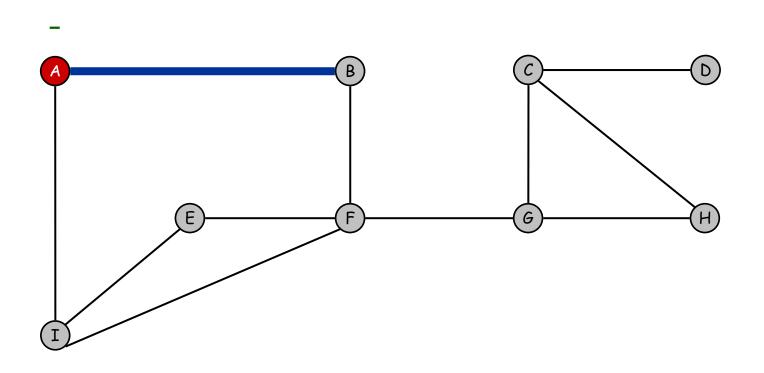
front

A



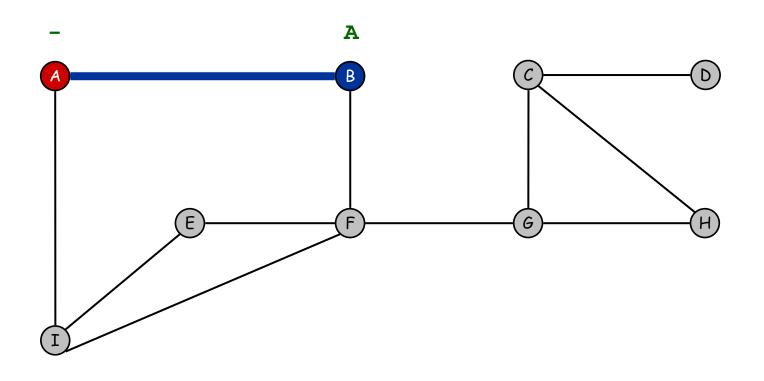
visit neighbors of A

front



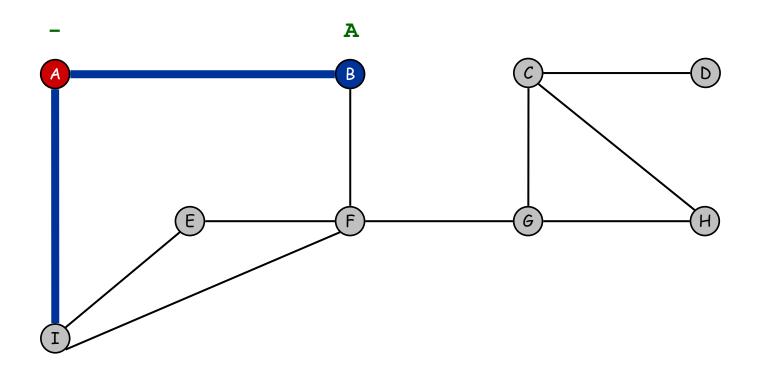
visit neighbors of A

front



B discovered front B

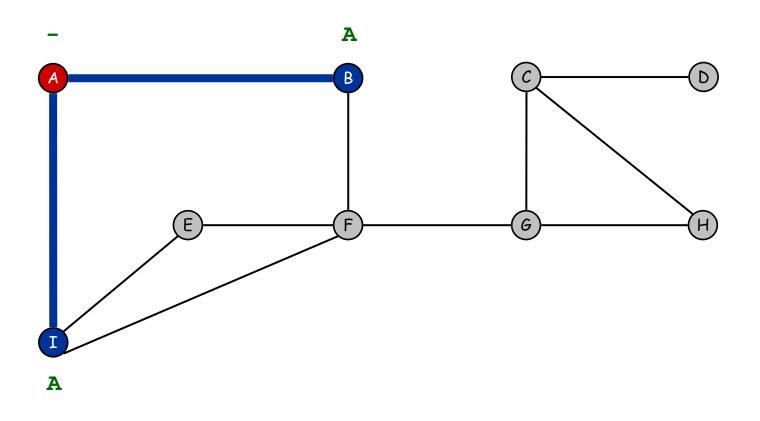
FIFO Queue



visit neighbors of A

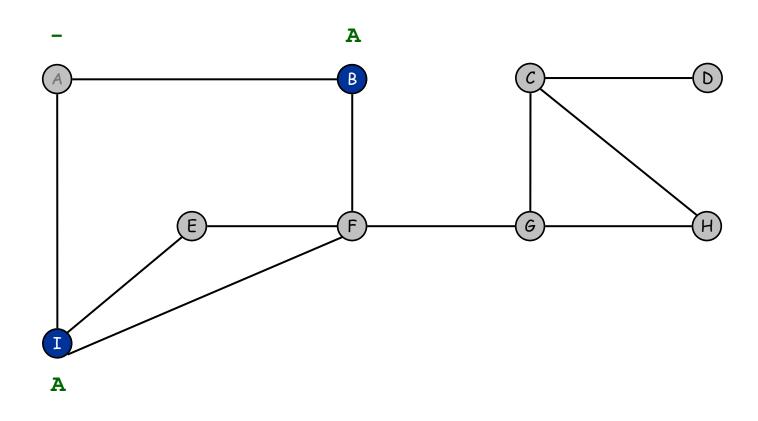
front

В



I discovered B I

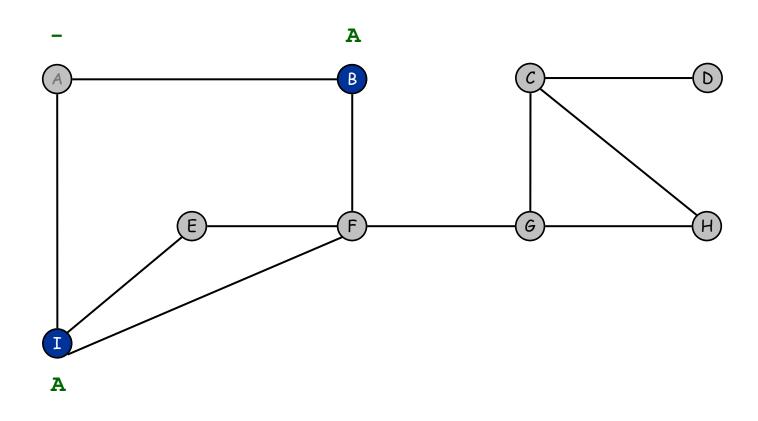
FIFO Queue



finished with A

front

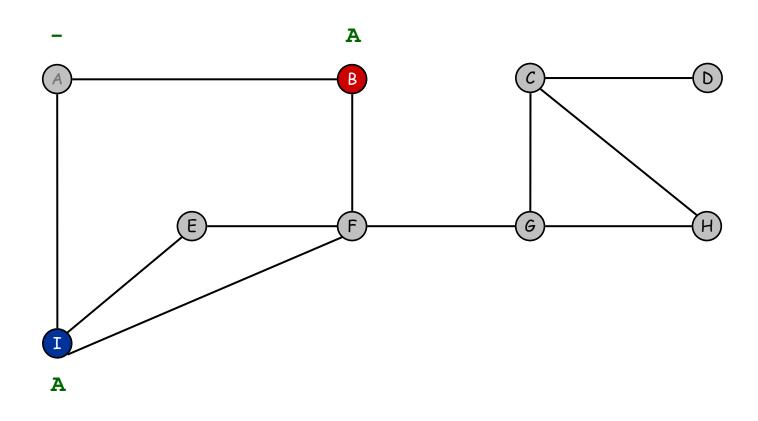
B I



dequeue next vertex

front

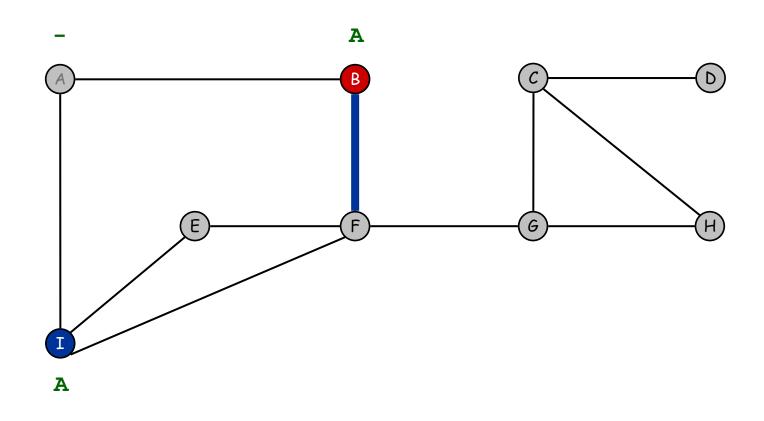
B I



visit neighbors of B

front

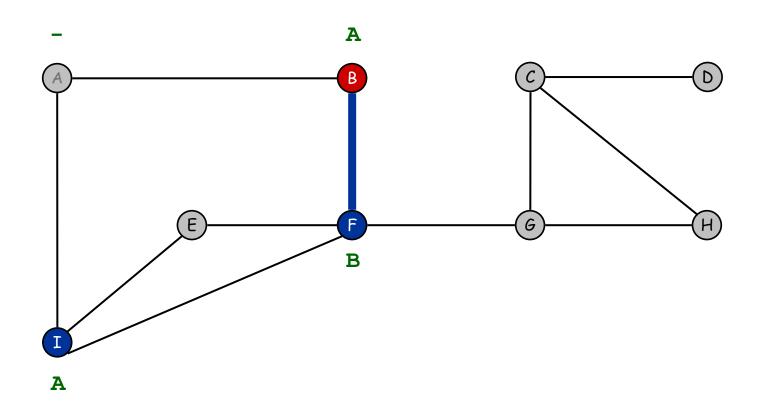
I



visit neighbors of B

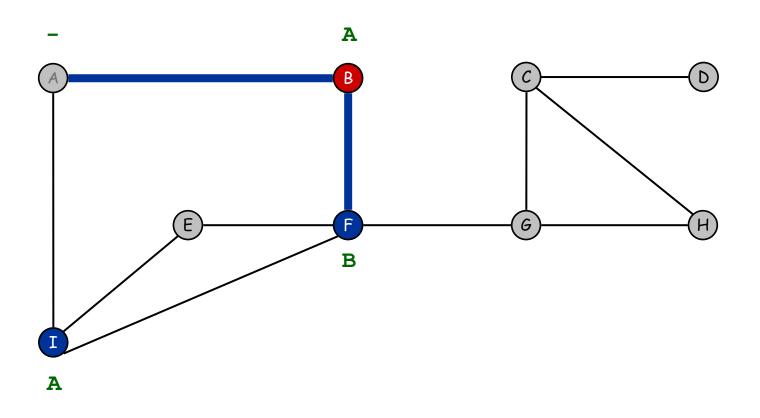
front

I



F discovered I F

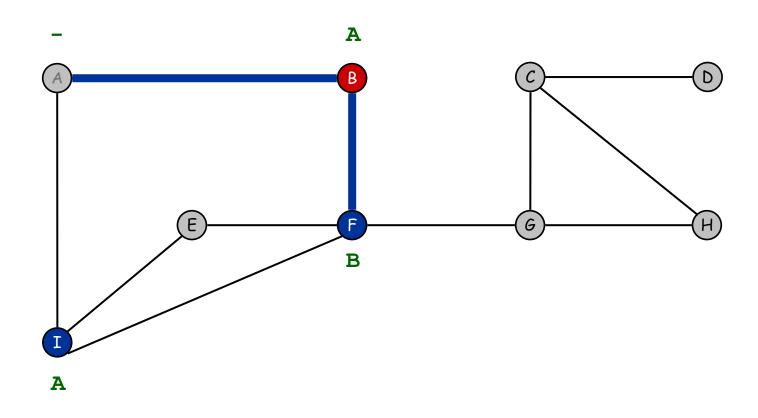
FIFO Queue



visit neighbors of B

front

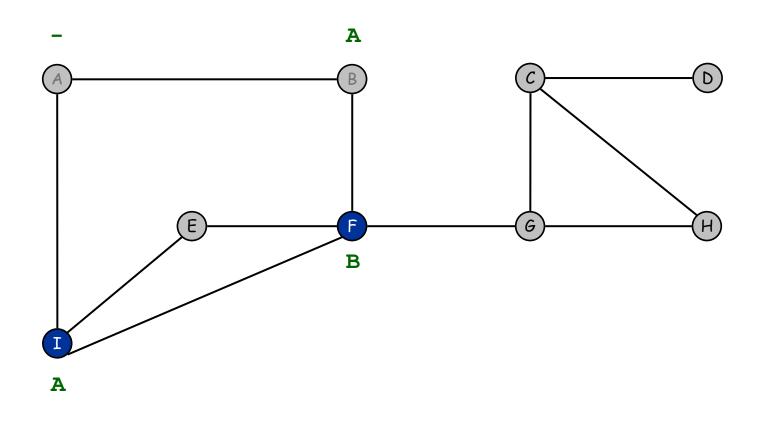
I F



A already discovered

front

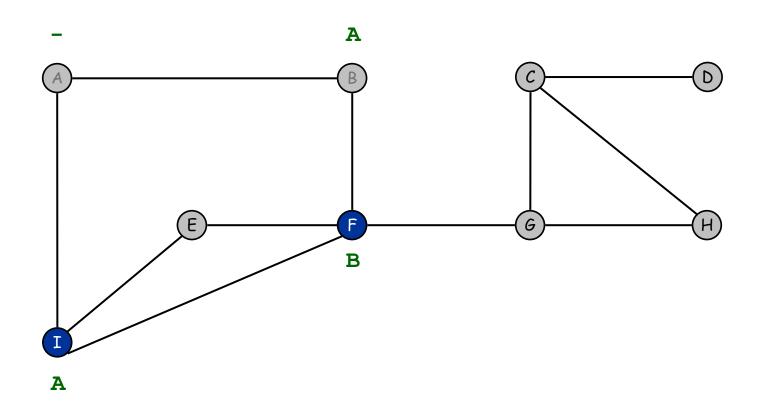
I F



finished with B

front

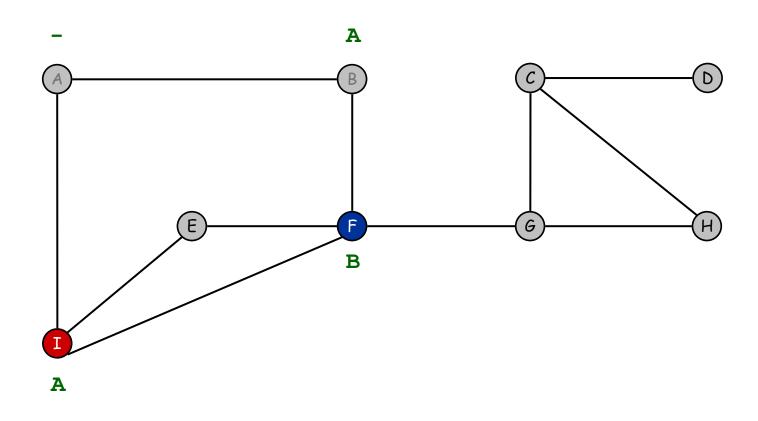
I F



dequeue next vertex

front

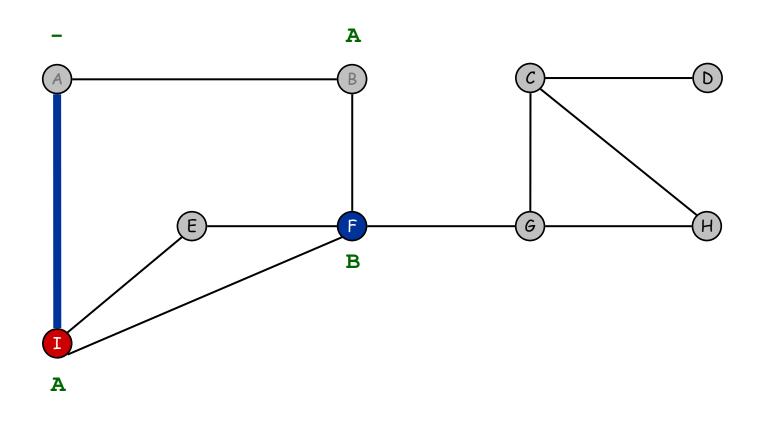
I F



visit neighbors of I

front

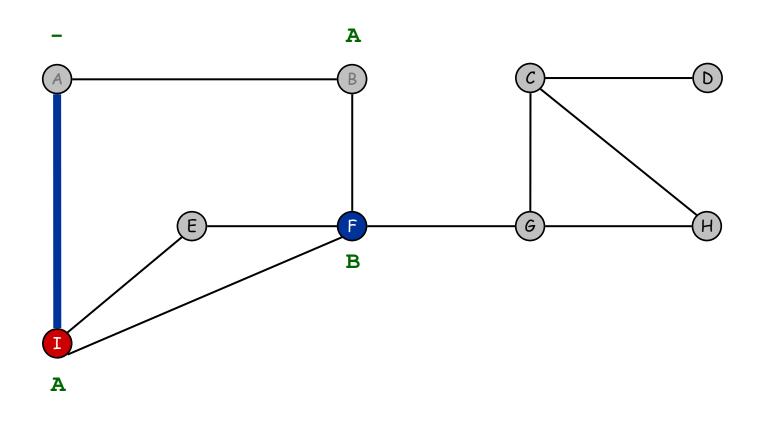
F



visit neighbors of I

front

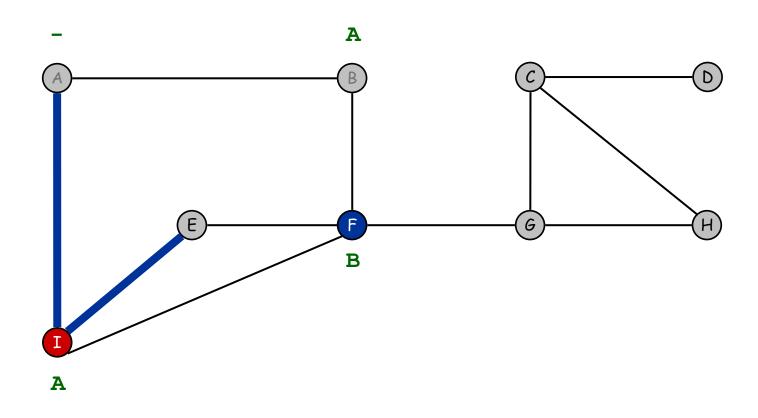
F



A already discovered

front

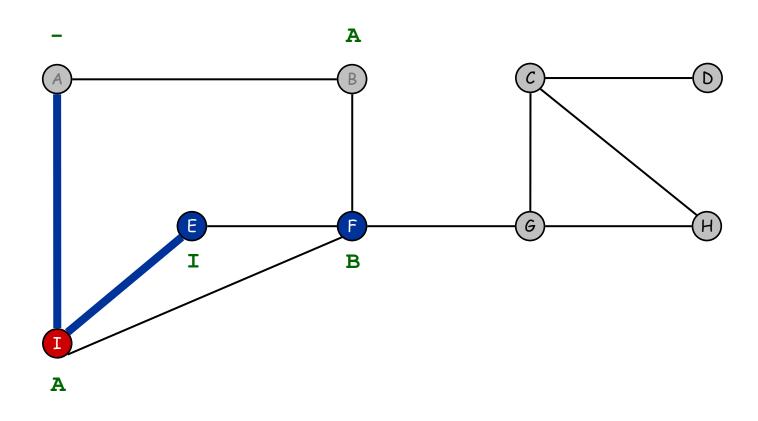
F



visit neighbors of I

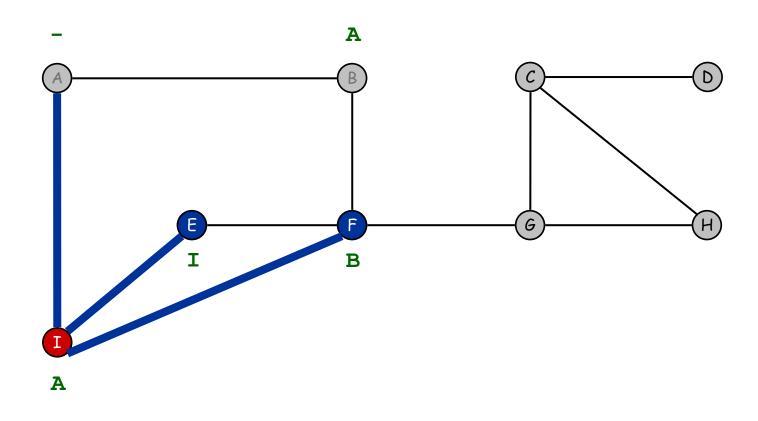
front

F



E discovered F E

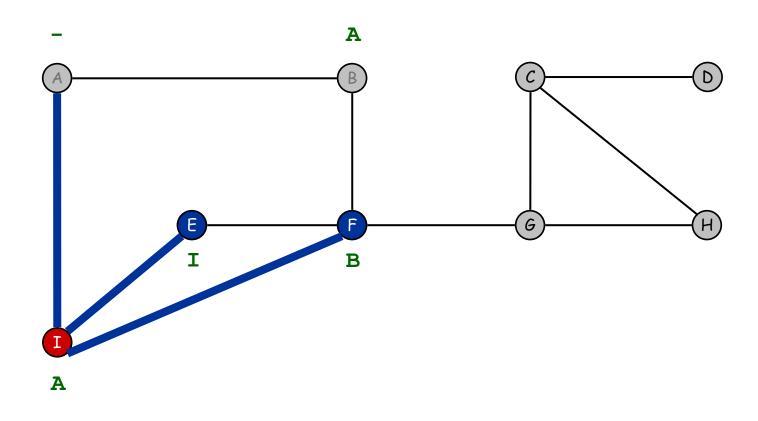
FIFO Queue



visit neighbors of I

front

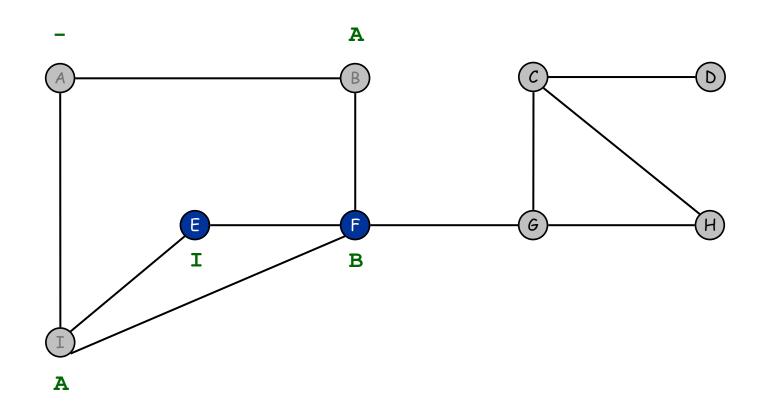
F E



F already discovered

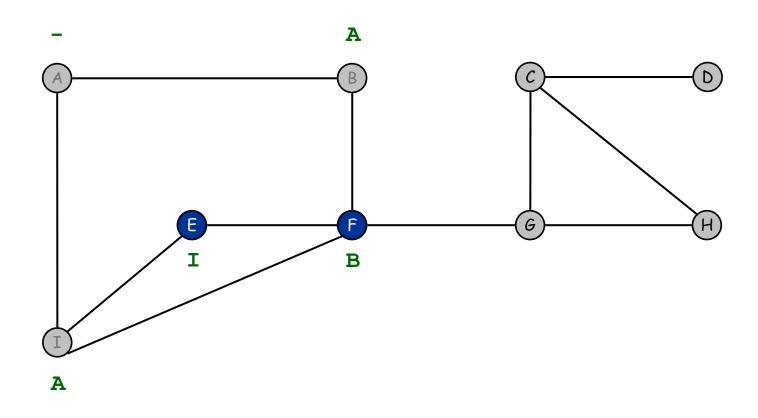
front

F E



I finished F E

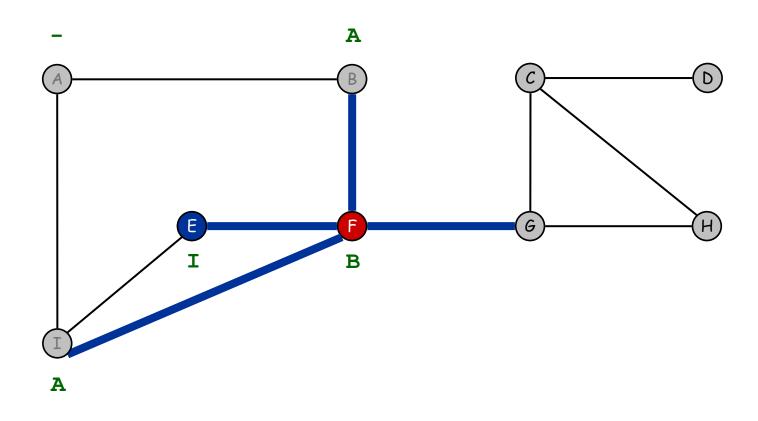
FIFO Queue



dequeue next vertex

front

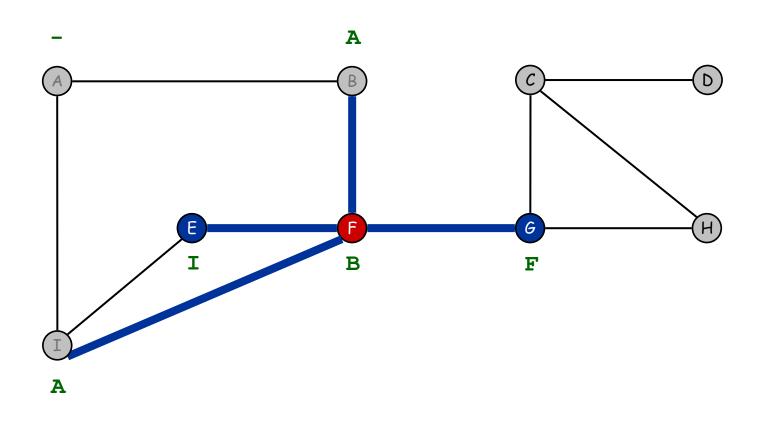
F E



visit neighbors of F

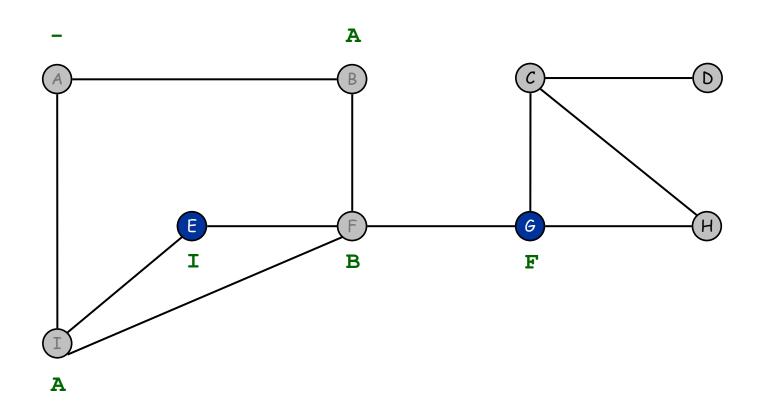
front

E



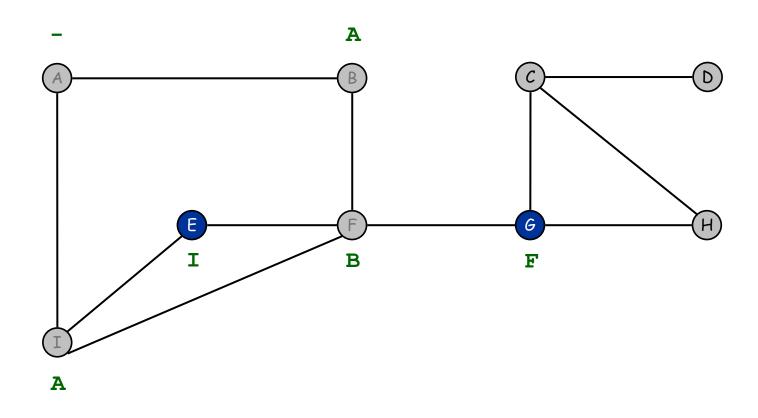
G discovered E G

FIFO Queue



F finished Front E G

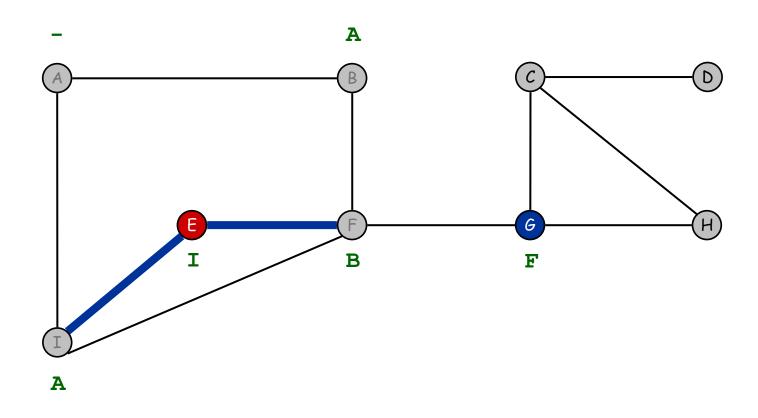
FIFO Queue



dequeue next vertex

front

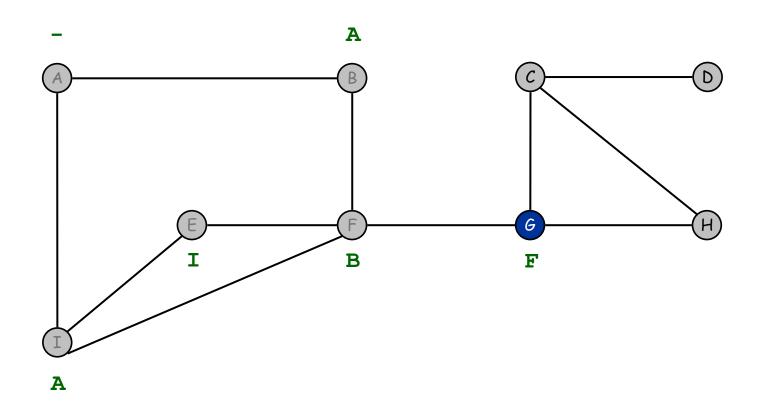
E G



visit neighbors of E

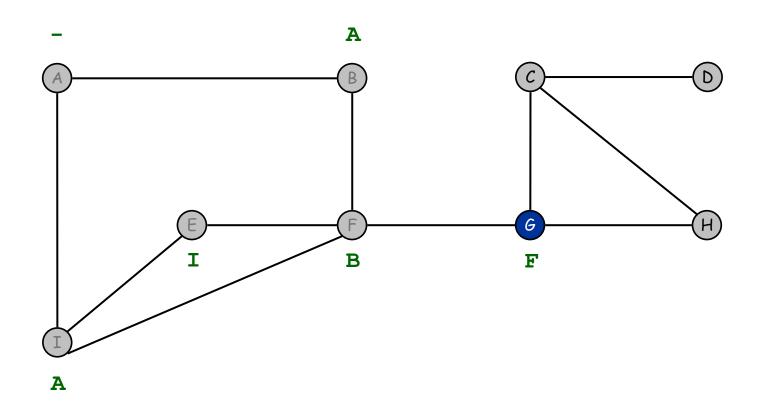
front

G



E finished G

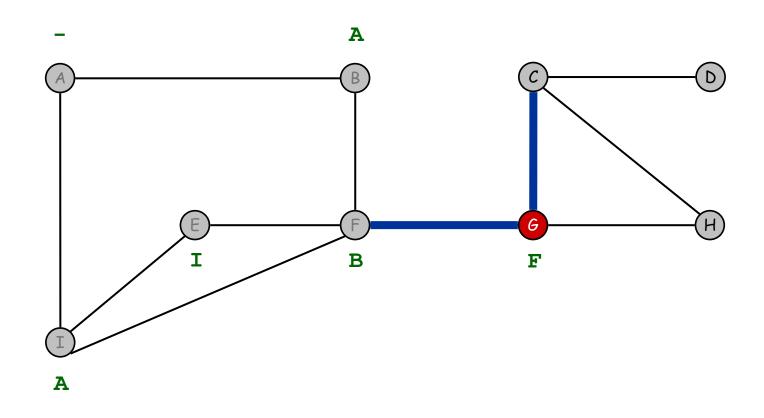
FIFO Queue



dequeue next vertex

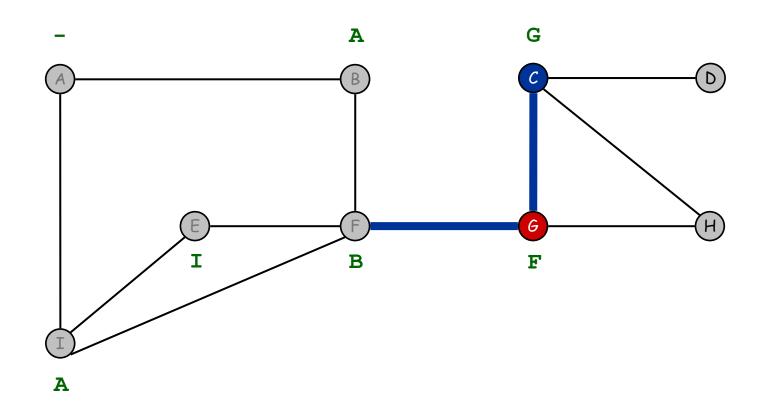
front

G

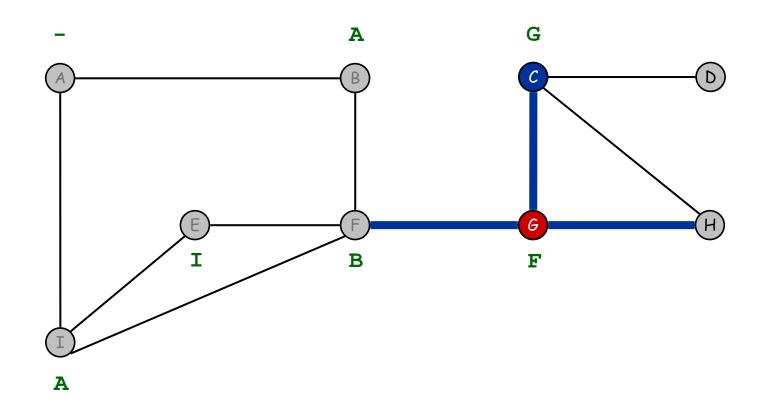


visit neighbors of G

front



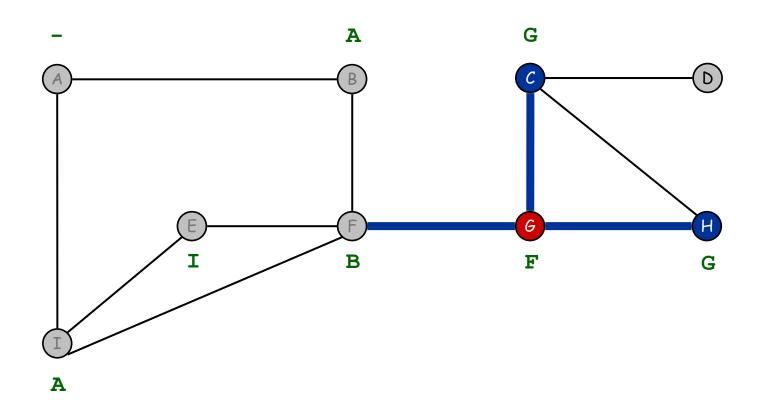
C discovered C



visit neighbors of G

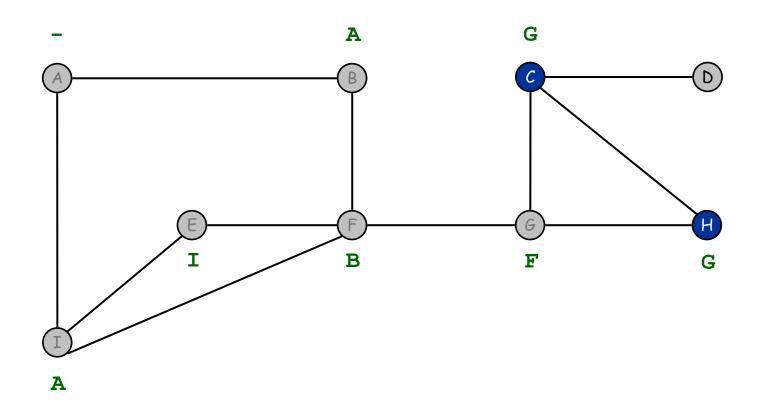
front

C



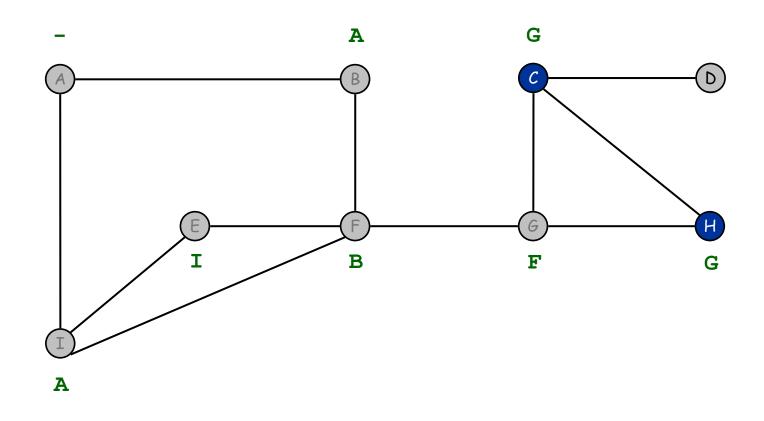
H discovered C H

FIFO Queue



G finished C H

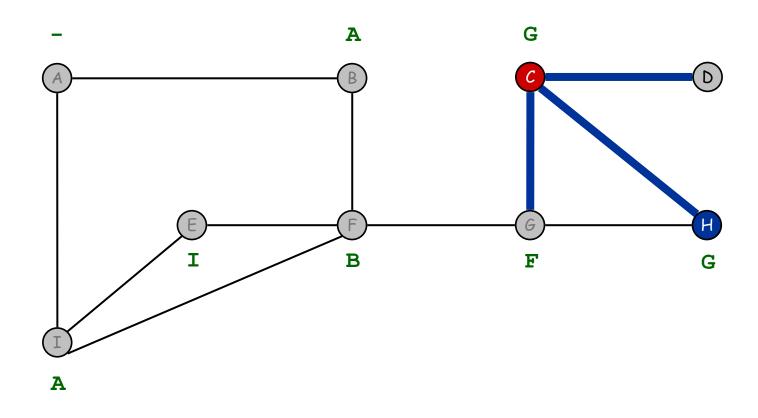
FIFO Queue



dequeue next vertex

front

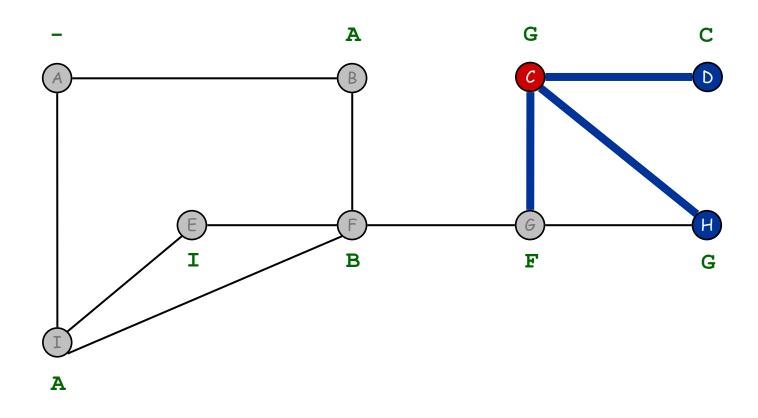
СН



visit neighbors of C

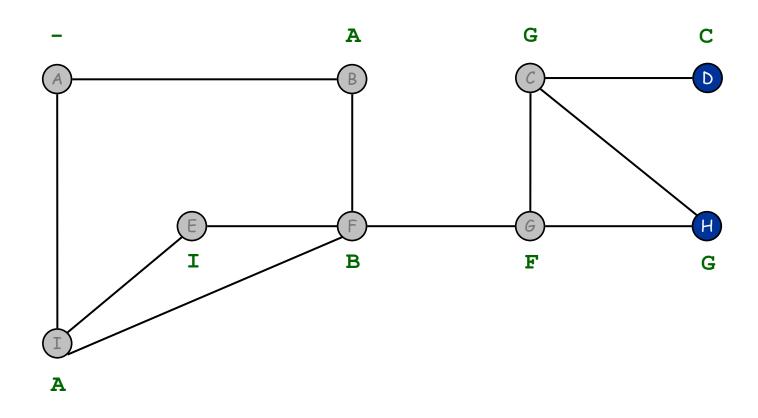
front

Н



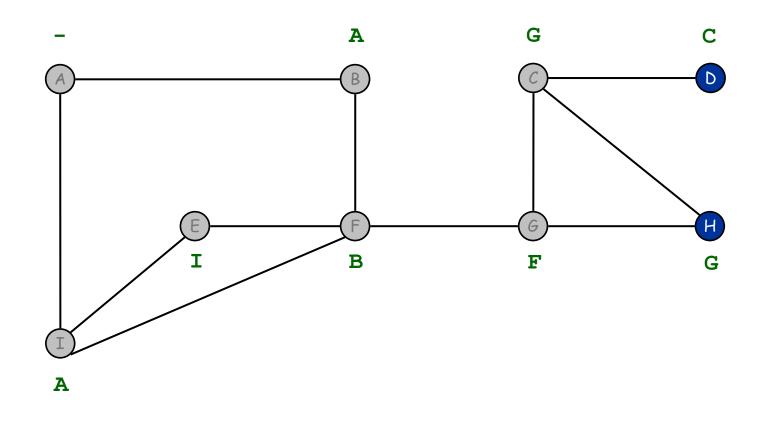
D discovered H D

FIFO Queue



C finished front H D

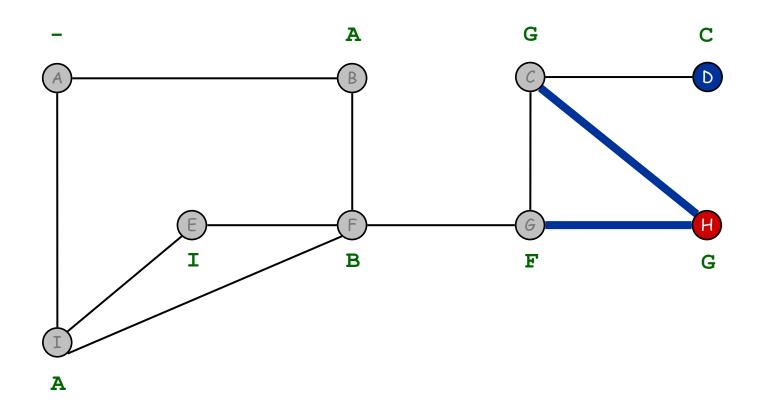
FIFO Queue



get next vertex

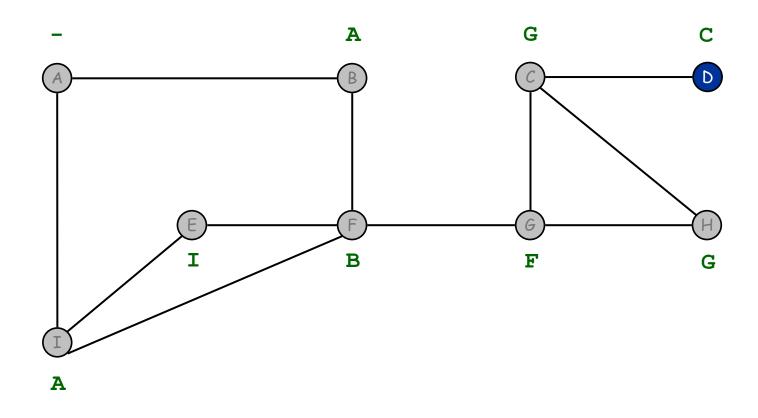
front

H D



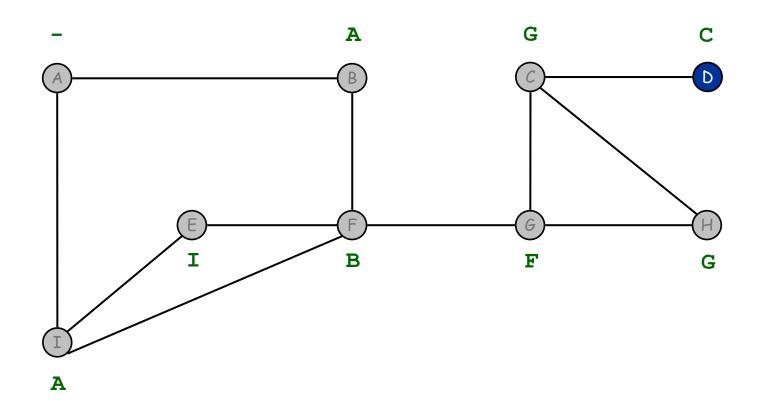
visit neighbors of H front D

FIFO Queue

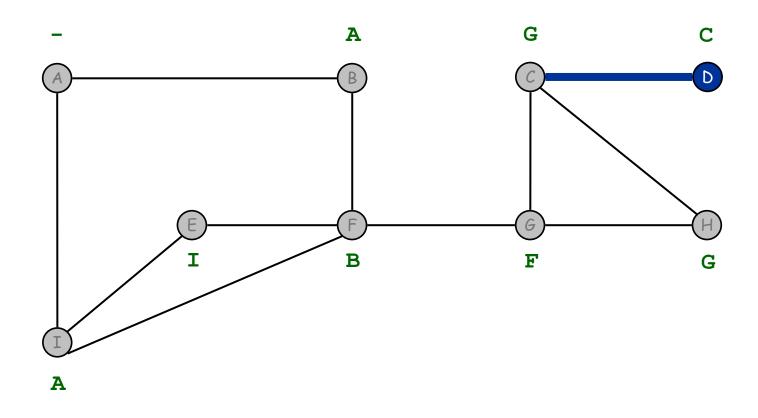


finished H front D

FIFO Queue

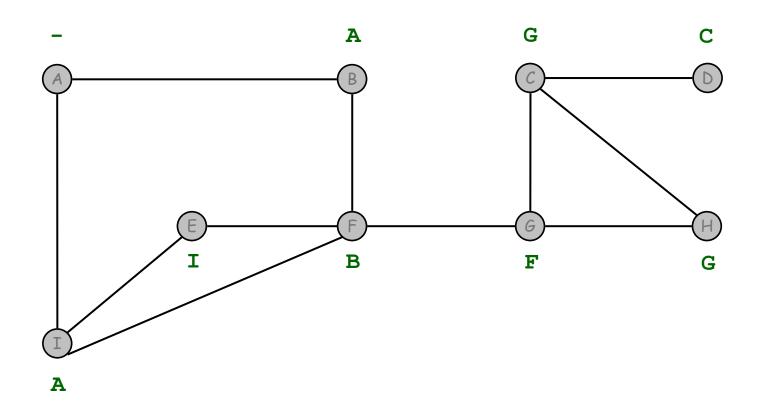


dequeue next vertex front D



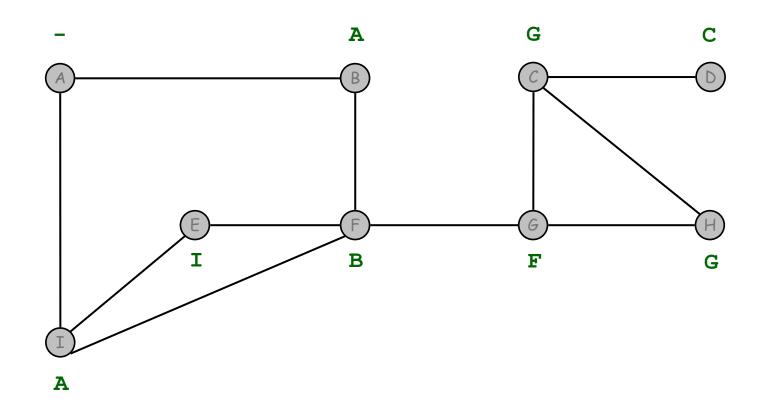
visit neighbors of D

front



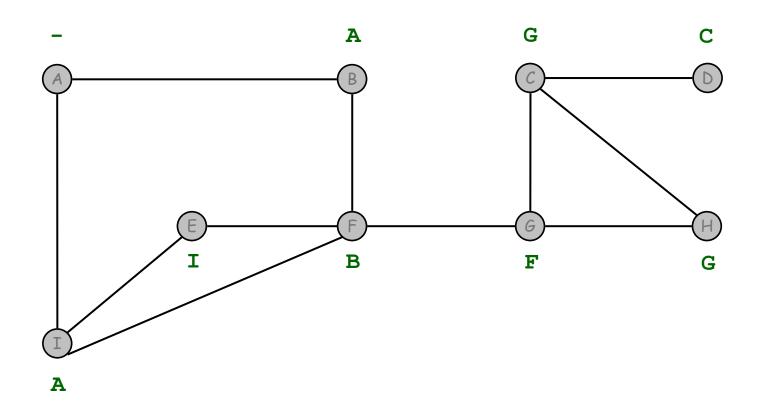
D finished

front



dequeue next vertex

front



STOP front

FIFO Queue

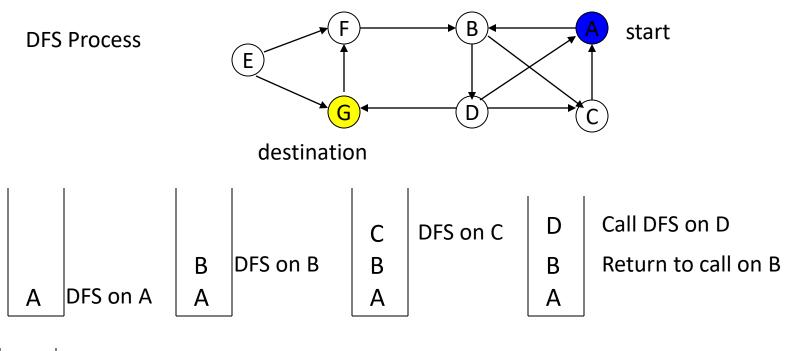
# Time Complexity

- Each visited vertex is put on (and so removed from) the queue exactly once
- When a vertex is removed from the queue, we examine its adjacent vertices
  - O(|V|) if adjacency matrix used
  - O(vertex degree) if adjacency lists used
- Total time
  - -O(|E||V|), where E is number of vertices in the component that is searched (adjacency matrix)= $O(|V|^2)$
  - O(|V| + sum of component vertex degrees) (adj. lists)
    - = O(|V| + number of edges in component)=O(|V|+|E|)

# Applications: Finding a Path

- Find path from source vertex s to destination vertex d
- Use graph search starting at s and terminating as soon as we reach d
  - Need to remember edges traversed
- Use depth first search ?
- Use breath first search?

# DFS vs. BFS

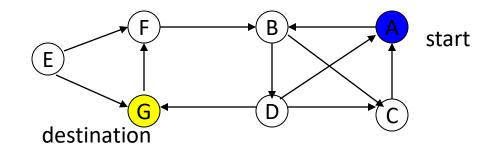


G Call DFS on G Found destination - done!
Path is implicitly stored in DFS recursion
Path is: A, B, D, G
B

Α

# DFS vs. BFS

**BFS Process** 



rear	front	
	A	
Initial call to BFS on A Add A to queue rear front		
	G	
Deque	eue D Add G	

rear	front	rear	front
	В		D C
Deque	eue A	Deque	eue B
A	Add B	Add	l C, D

D	
Dequeue C	
Nothing to add	

rear

front

found destination - done! Path must be stored separately

## Path From Vertex s To Vertex d

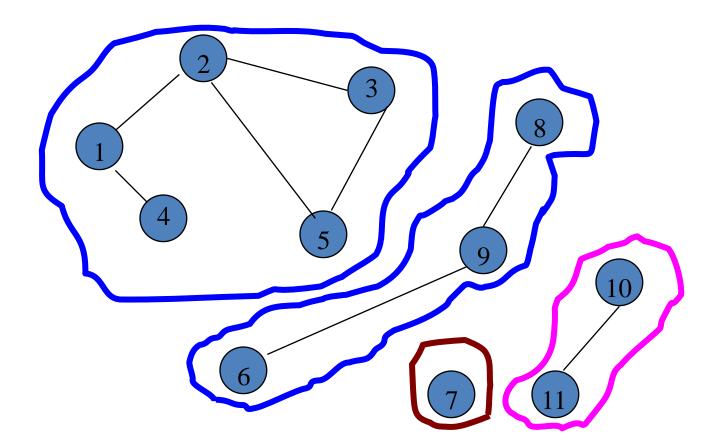
- Time
  - $-O(|V|^2)$  when adjacency matrix used
  - -O(|V|+|E|) when adjacency lists used (|E| is number of edges)

# Is The Graph Connected?

- Start a breadth-first search at any vertex of the graph
- Graph is connected iff all n vertices get visited
- Time
  - $^{\bullet}$  O( $|V|^2$ ) when adjacency matrix used
  - O(|V|+|E|) when adjacency lists used (|E| is number of edges)

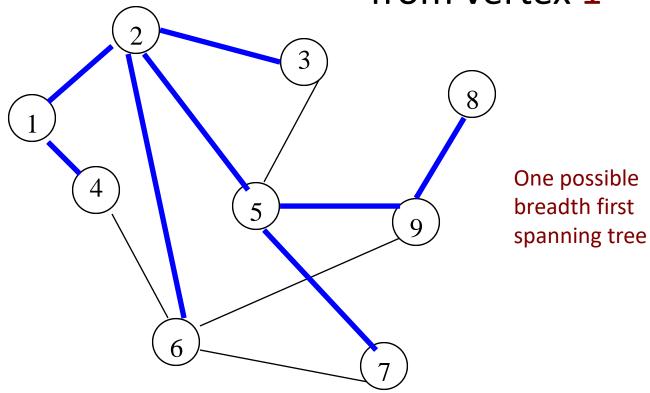
# **Connected Components**

- Start a BFS at any as yet unvisited vertex of the graph
- Newly visited vertices (plus edges between them) define a component
- Repeat until all vertices are visited



# Breadth First Spanning Tree

Breadth-first search from vertex 1



- Keep track of edges used to reach new vertices
- These edges form a spanning tree if the graph is connected

# Spanning Tree

- Start a breadth-first search at any vertex of the graph
- If graph is connected, the n-1 edges used to get to unvisited vertices define a spanning tree (breadth-first spanning tree)
- Time
  - $-O(V^2)$  when adjacency matrix used
  - -O(V+E) when adjacency lists used (E is number of edges)