Data Structures and Advanced Programming

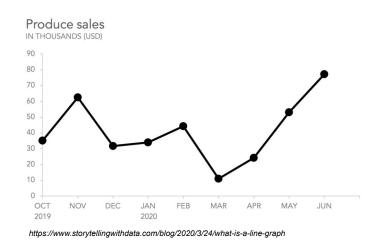
Fu-Yin Cherng National Taiwan University

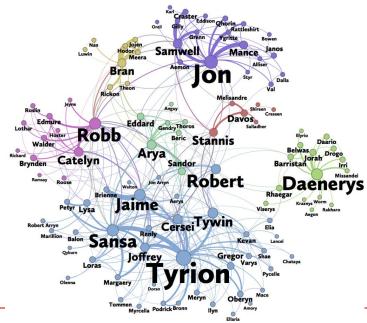
Graph

What is graph?

a way to illustrate data & represent the relationships among

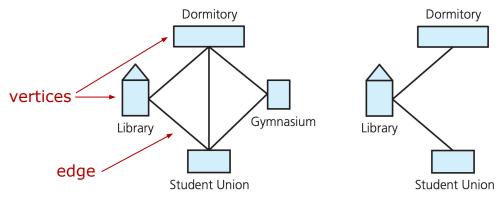
data items



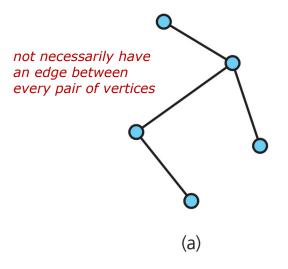


What is graph?

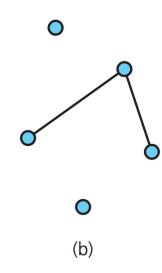
- General definition of graph G consists of 2 sets
 - V: a set of vertices/nodes
 - **E**: a set of edges that connect the vertices
- Subgraph: a subset of a graph vertices and its edges



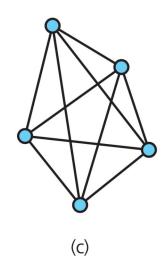
- Adjacent vertices
 - 2 vertices are adjacent if joined by an edge
- Path
 - a path between two vertices is a sequence of edges the begins at one vertex and ends at another vertex
- Simple Path
 - a path passes through a vertex only once
- Cycle
 - a path that begins and ends at the same vertex



Connected Graph each pair of distinct vertices has a path between them

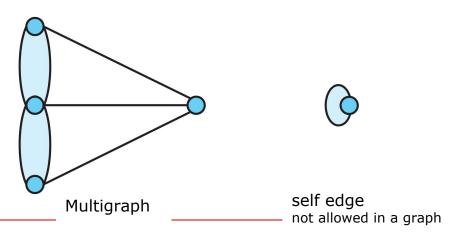


Disconnected Graph

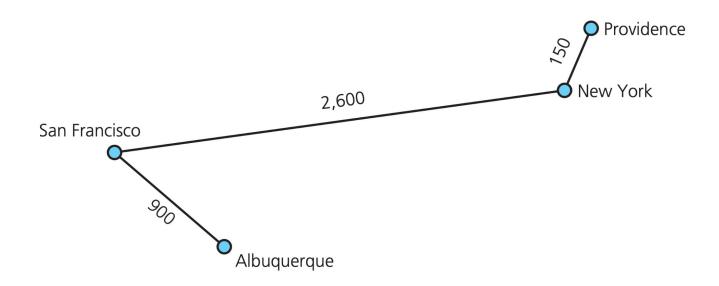


Complete Graph each pair of distinct vertices has an edge between them

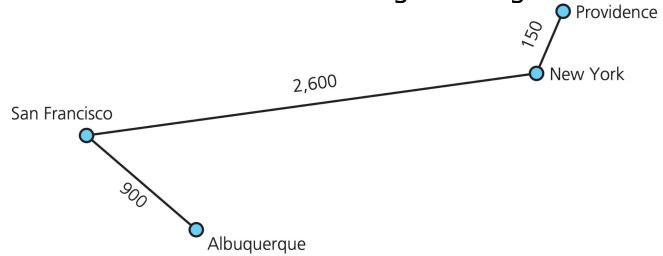
- A graph
 - edge cannot begin and end at the same vertex (self edge)
 - cannot have duplicate edges between vertices
- Multigraph is not a graph and can have multiple edges



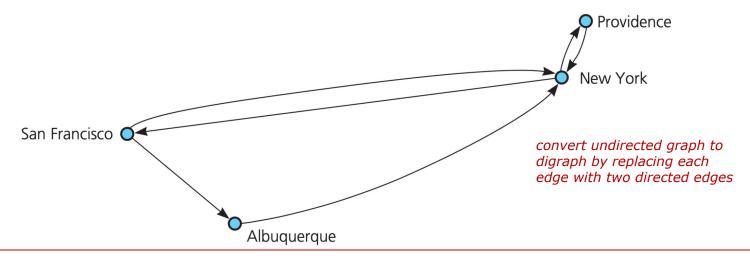
Weighted graph: edges with numeric labels



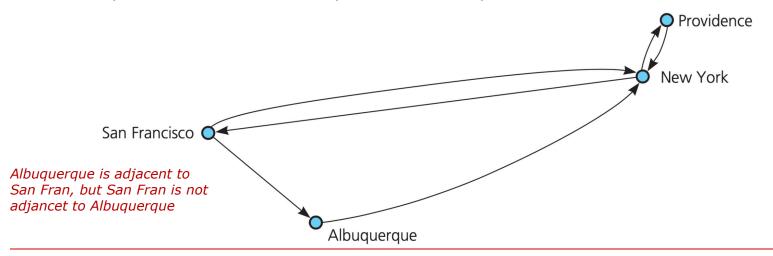
 Undirected graphs: edges do not indicate direaction, can travel either direction along the edges



- Directed graph or digraph: graph with directed edge
 - can have 2 edges with different directions between a pair of vertices



- Directed graph or digraph
 - adjacent vertices: if there is a directed edge from vertex x to vertex y, then y is adjacent to x (x -> y)
 - y is successor of x, x is predecessor of y



Abstract data type: Graph

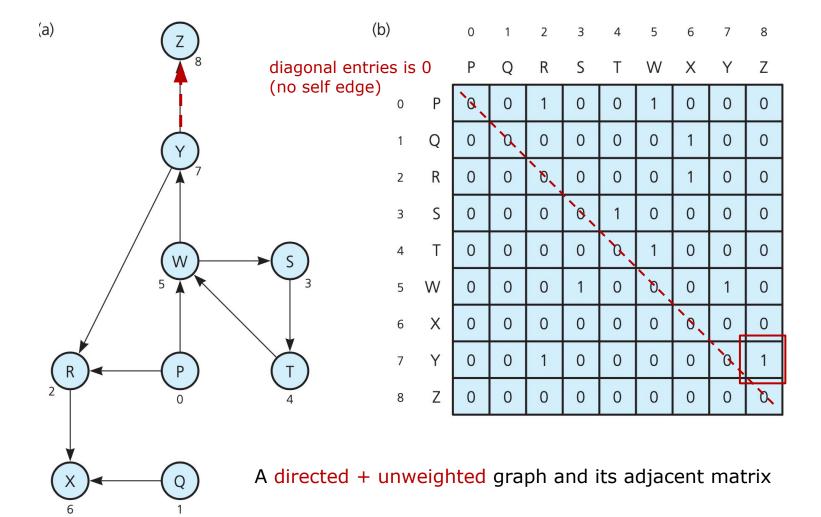
- define following ADT graph's vertices contain value
 - you can also design another ADT graph whose vertices don't contain value.
- many operations, for example
 - Test whether a graph is empty.
 - Get the number of vertices in a graph.
 - Insert an edge between two given vertices in a graph
 - Retrieve from a graph the vertex that contains a given value
 - **...**

Abstract data type: Graph

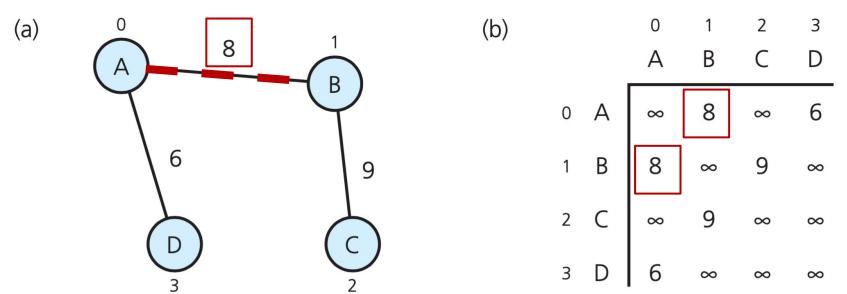
```
//An interface for the ADT undirected, connected graph
template < class LabelType >
class GraphInterface{
    public:
          virtual int getNumVertices() const = 0;
          virtual int getNumEdges() const = 0;
          virtual bool add(LabelType start, LabelType end, int edgeWeight) = 0;
          irtual bool remove(LabelType start, LabelType end) = 0;
          virtual int getEdgeWeight(LabelType start, LabelType end) const = 0;
          virtual void depthFirstTraversal(LabelType start, void visit(LabelType&)) = 0;
          virtual void breadthFirstTraversal(LabelType start, void visit(LabelType&)) =
    ():
```

Implementing Graphs: adjacency matrix

- use adjacency matrix or adjaceny list
- $_{\Box}$ for a graph with *n* vertices numberred 0, 1, ..., n-1
 - adjacent matrix: matrix[n][n]
 - unweighted:
 - matrix[i] [j] is 1 (true) if there is an edge from vertex i
 to vertex j, and 0 (false) otherwise
 - weighted
 - matrix[i][j] is the weight that labels the edge from vertex i to vertex j, and ∞ otherwise



undirected graph is symmertical



A undirected + weighted graph and its adjancet matrix

Implementing Graphs: adjacency matrix

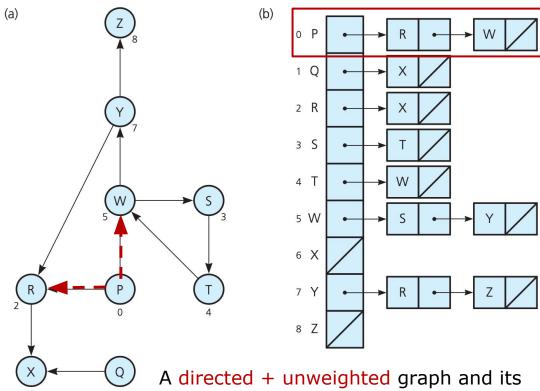
- use second array to represent the n vertex values
- values[i] is the value in vertex i

Implementing Graphs: adjacency list

- use adjacency matrix or adjacency list
- \Box for a graph with *n* vertices numberred 0, 1, ..., n-1
 - adjancey list: consists of n linked chains
 - see example directly

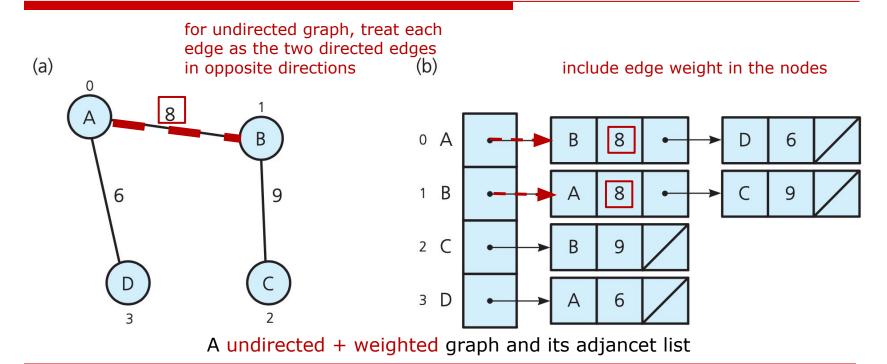
Implementing Graphs: adjacency list

- the ith linked chain has nodes for vertices with edge to vertex i
- If the vertex has no value, the node needs to contain some indication of the vertex's identity



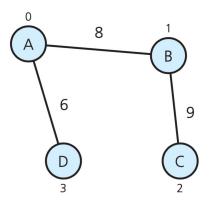
adjancet list

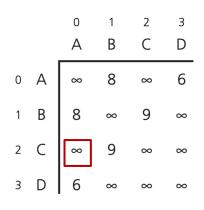
Implementing Graphs: adjacency list

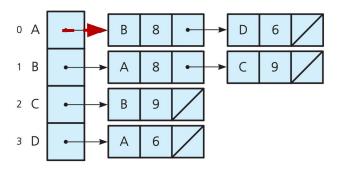


- Which one is better? adjacency matrix or list?
- depend on applications and operations

- For example, adjacency matrix is more efficient for determining whether there is an edge from vertex i to vertex j
 - only need to check the value of matrix[i][j]
 - if an adjaceny list, need to traverse ith linked chain



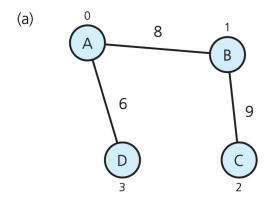


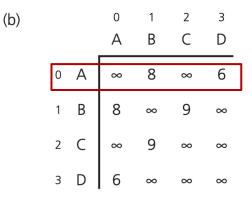


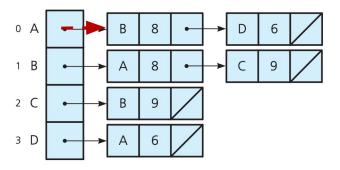
edge between A and C? _____ check value in [A][C] or [C][A]

_ traverse 2th linked chain to check if there is node for C

- On the other hand, adjacency list is more efficient for finding all vertices adjacent to a given vertex i
 - directly traverse the ith linked chain.





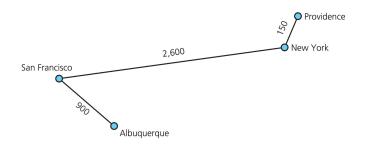


find adjacency of A _____ traverse all columns in 0th row

____ traverse 0th linked chain with 2 nodes

- storage space requirement of a directed graph with n node
 - adjacency matrix always has n*n entries
 - adjacency list has number of nodes equals the number of edges in a directed graph
- adjacency list often requires less storage than an adjacency matrix.

- choosing a graph implementation for a particular application
 - what operations you will perform most frequently
- for example, the flight map problem, most frequent operation was to find all cities (vertices) adjacent to given city
 - adjacency list would be more efficient



Graph Traversal

- traverse all vertices that it can reach
- mark each vertex during visit and only visit a vertex once
 - to prevent loop indefinitely due to a cycle in a graph
- 2 basic graph-traversal algorithms
 - Depth-First Search (DFS)
 - Breadth-First Search (BFS)

Depth-First Search (DFS)

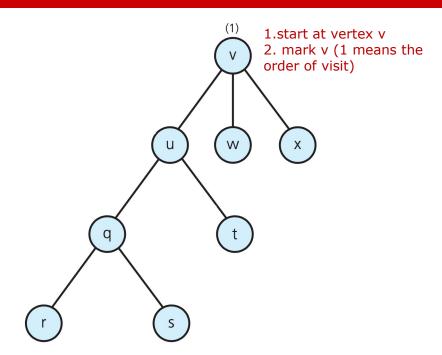
- □ 深度優先
- From vertex v, DFS traversal goes as far (deep) as possible from the vertex v before backing up
- After visiting a vertex, a DFS vists an unvisited adjacent vertex if possible.

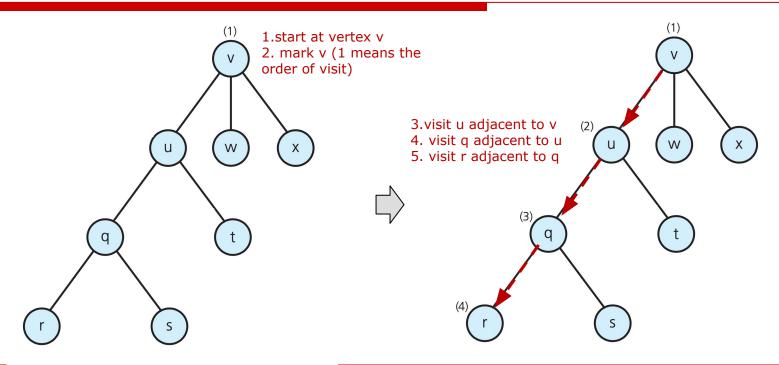
Depth-First Search (DFS)

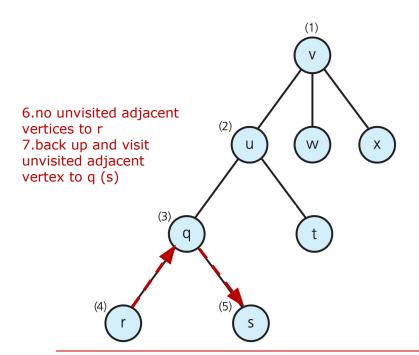
DFS strategy has a recursive form

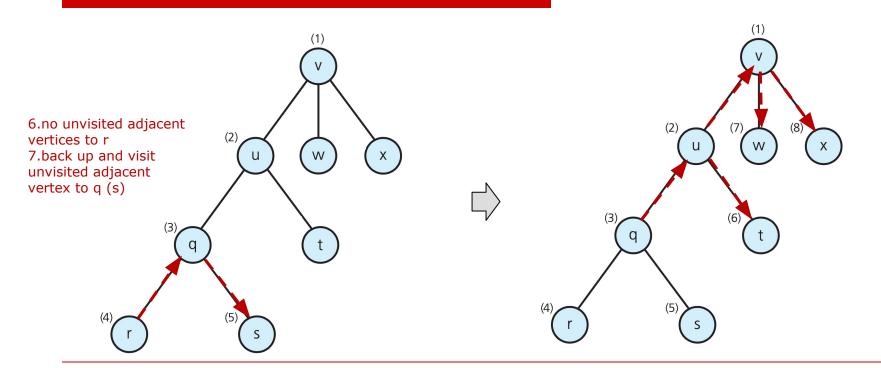
```
// Traverses a graph beginning at vertex v by using a
// depth-first search: Recursive version.
dfs(v: Vertex)
   Mark v as visited
   for (each unvisited vertex u adjacent to v)
        dfs(u)
```

- don't specify the order of which it should visit the vertices adjacent to v
- Could visit the adjacent vertices in sorted order alphabetically or numerically





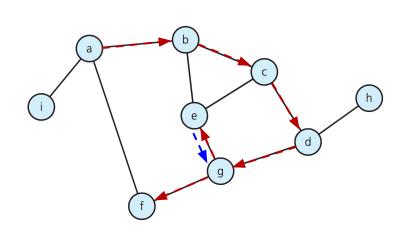




Depth-First Search (DFS)

DFS strategy has a iterative version using a stack

```
// Traverses a graph beginning at vertex v by using a
// depth-first search: Iterative version.
dfs(v: Vertex)
     s = a new empty stack
    s.push(v)
    Mark v as visited
    while (!s.isEmpty()) {
          if (no unvisited vertices are adjacent to the vertex on the top of the stack)
               s.pop() // Backtrack
          else{
               Select an unvisited vertex u adjacent to vertex on the top of the stack
               s.push(u) // vertex u on the top of the stack
               Mark u as visited
```



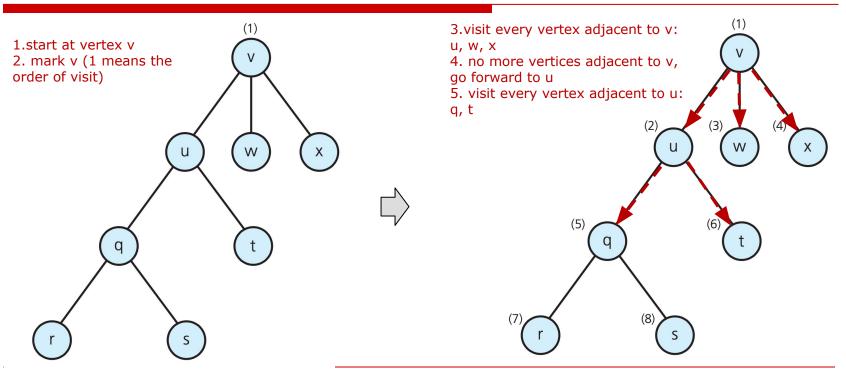
order of visit: a, b, c, d, e, f, h, i

Node visited Stack (bottom to top) a b abc a b c d abcdg a b c d g e no unvisited vertex adjacent to e (backtrack) a b c d g pop e, so the top vertex in stack is g abcdgf visit f adjacent to g (backtrack) a b c d g pop f, so the top vertex is g a b c d pop g, so the top vertex is d (backtrack) abcdh (backtrack) a b c d (backtrack) a b c (backtrack) a b (backtrack) a i (backtrack) (backtrack) (empty)

Breadth-First Search (BFS)

- □ 廣度優先
- BFS traversal visits all vertices adjacent to a vertex before going forward (deeper)

Breadth-First Search (BFS): example



Breadth-First Search (BFS)

- BFS strategy has a recursion version using a queue
- not as simple as the recursive version of DFS traversal

```
// Traverses a graph beginning at vertex v by using a
// breadth-first search: Recursive version.
bfs(q: Queue)
   if (q.empty()) return

v = q.dequeue(the front node)

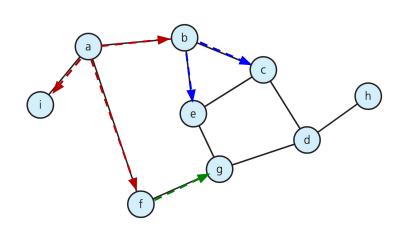
for(each unvisited vertex u adjacent to v){ // do for every edge `v -> u
        Mark u as visited
        q.enqueue(u)
}
bfs(q)
```

Breadth-First Search (BFS)

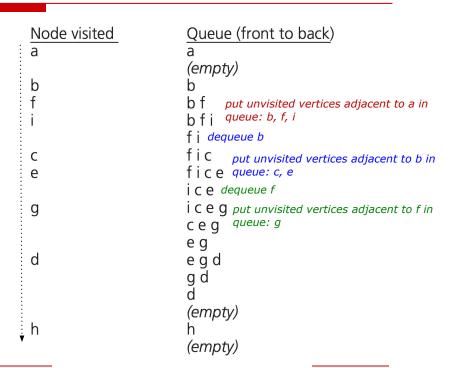
BFS strategy has a iterative version using a queue

```
// Traverses a graph beginning at vertex v by using a
// breadth-first search: Iterative version.
bfs(v: Vertex)
     q = a new empty queue
     q.enqueue(v) // Add v to queue and mark it
     Mark v as visited
     while (!q.isEmpty()) {
          q.dequeue (the front node)
          for (each unvisited vertex u adjacent to v) { // do for every edge v -> u
               Mark u as visited
               q.enqueue(u)
```

Breadth-First Search (BFS): example 2



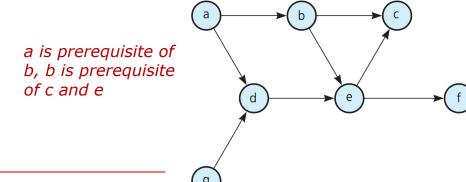
order of visit: a, b, f, i, c, e, g, d, h



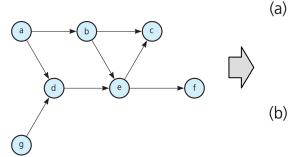
Application of Graphs

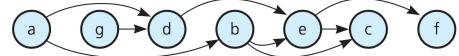
- Some common applications of graphs
 - Topological Sorting
 - Shortest Paths

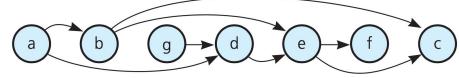
- A directed graph without cycle has a natural order
- If the vertices represent courses, the graph represent the prerequisite structure for the courses
- In what order should you take all seven courses so that you will satisfy all prerequisites? Topological order



- The vertices in a given graph may have several topological orders
 - a, g, d, b, e, c, f
 - a, b, g, d, e, f, c

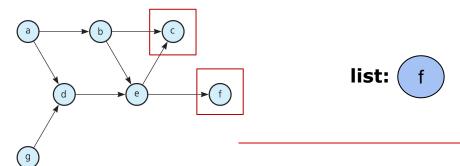




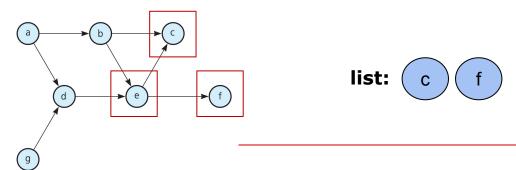


arrange the vertices of a directed graph in a topological order, edges will all point in one direction

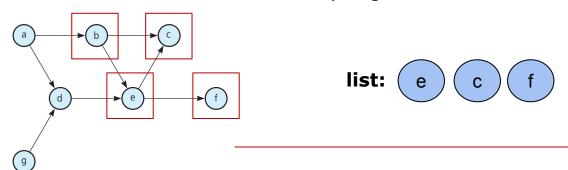
- Arranging the vertices into a topological order is called topological sorting
- Several ways to do topolgical sorting
 - find a vertex that has no successor
 - remove the vertex and all edges lead to it from graph
 - add the vertex to the beginning of the list
 - repeatedly the above steps until the graph is empty
 - the list of vertices is in topological order



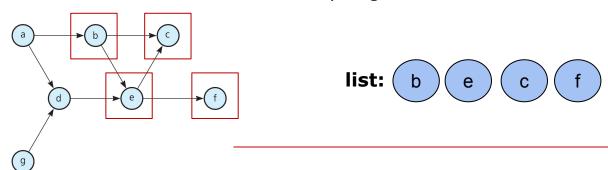
- Arranging the vertices into a topological order is called topological sorting
- Several ways to do topolgical sorting
 - find a vertex that has no successor
 - remove the vertex and all edges lead to it from graph
 - add the vertex to the beginning of the list
 - repeatedly the above steps until the graph is empty
 - the list of vertices is in topological order



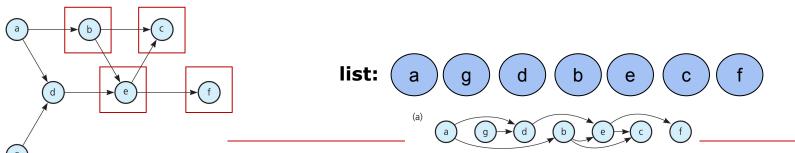
- Arranging the vertices into a topological order is called topological sorting
- Several ways to do topolgical sorting
 - find a vertex that has no successor
 - remove the vertex and all edges lead to it from graph
 - add the vertex to the beginning of the list
 - repeatedly the above steps until the graph is empty
 - the list of vertices is in topological order



- Arranging the vertices into a topological order is called topological sorting
- Several ways to do topolgical sorting
 - find a vertex that has no successor
 - remove the vertex and all edges lead to it from graph
 - add the vertex to the beginning of the list
 - repeatedly the above steps until the graph is empty
 - the list of vertices is in topological order



- Arranging the vertices into a topological order is called topological sorting
- Several ways to do topolgical sorting
 - find a vertex that has no successor
 - remove the vertex and all edges lead to it from graph
 - add the vertex to the beginning of the list
 - repeatedly the above steps until the graph is empty
 - the list of vertices is in topological order



Topological Sorting: pseudocode

```
// Arranges the vertices in graph the Graph into a
// topological order and places them in list aList
topSort1(theGraph: Graph, aList: List)
    n = number of vertices in theGraph
    for (step = 1 through n) { //all vertices
        Select a vertex v that has no successors
        aList.insert(1, v) //insert at the beginning of list
        Remove from the Graph vertex v and its edges
```

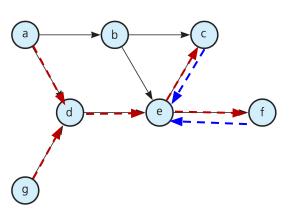
Topological Sorting: DFS topological sorting

- depth-first search (DFS) topological sorting
 - push all vertices that have no predecessor onto a stack
 - Each time you pop a vertex from the stack
 - add it to the beginning of a list of vertices
- modified from DFS traversal

Topological Sorting: DFS pseudocode

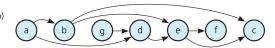
```
topSort2(theGraph: Graph, aList: List)
     s = a new empty stack
     for (all vertices v in the graph) {
           if (v has no predecessors) {
                s.push(v)
                Mark v as visited
     while(!s.isEmpty()){
           if (all vertices adjacent to the vertex on the top of the stack have been visited)
                s.pop(v) //backtrace
                aList.insert(1, v)
           else{
                Select an unvisited vertex u adjacent to the vertex on the top of the stack
                s.push(u)
                Mark u as visited
```

Topological Sorting: DFS



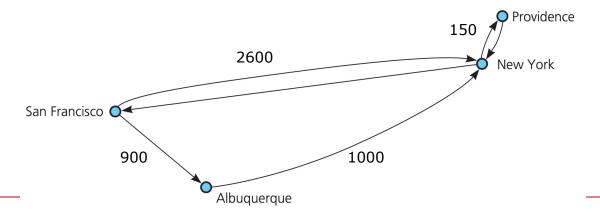
Action	Stack s (bottom to top)	ist aList (beginning to end)
Push a		ssors, g at the top of stack
Push g Push d	<pre>a g a g d Select an unvisite</pre>	d vertex d adjacent to g
Push e	agde Select an unvisite	d vertex e adjacent to d
Push c	agdec Select an unvisit	ted vertex c adjacent to e
Pop c, add c to aList	a g d e	cc has no adjacent & unvisited vertex
Push f	a g d e f	С
Popf, add f to aList	a g d e	\mathfrak{f}_{C} add f at the biginning of list
Pop e, add e to aList	a g d	e f c
Pop d, add d to aList	a g	d e f c
Pop g, add g to aList	a	g d e f c
Push b	a b	g d e f c
Pop b, add b to aList	a	b g d e f c
Pop a, add a to aList	(empty)	a b g d e f c

resulting topological order: (b)



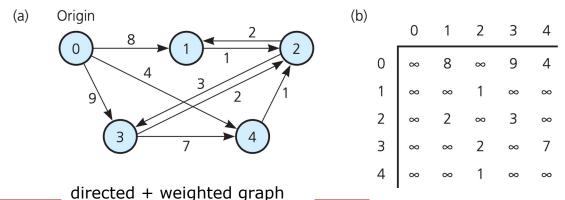
Shortest Path

- find the shortest path between 2 cities (vertices)
 - the path has the smallest edge-weight sum
- the weight can also be the dollars or duration (nonnegative)
- the sum of weights of edges of a path called path's length/weight/cost



Shortest Paths: example

- shortest path between vertex 0 to 1
 - not edge between 0 and 1 (cost 8)
 - path: 0 to 4 to 2 to 1 (cost 7)
- □ origin labeled 0, other vertices labeled from 1 to n-1



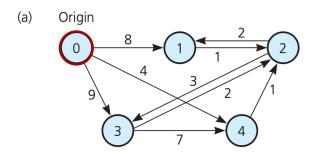
matrix of the left graph

- Dijkstra's shortest-path algorithm
 - determines the shortest paths between a given origin and all other vertices
- vertexSet
 - a set of selected vertices
- weight
 - array where weight [v] is the weight of shortest path from vertex 0 (origin) to vertex v that passes through vertices in vertexSet

Initially,

- vertexSet contains only vertex 0 (origin), and weight contains the weights of the single-edge paths from vertex 0 to all other vertices v
- weights[v] = matrix[0][v]

initial weight is the first row of adjacency matrix

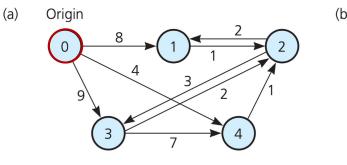


(b)					•		
(D)		0	1	2	3	4	
	0	8 8 8 8	8	∞	9	4]
	1	∞	∞	1	∞	∞	_
	2	∞	2	∞	3	∞	
	3	∞	∞	2	∞	7	
	4	∞	∞	1	∞	∞	

- After this initialization step
 - find a vertex v
 - □ **not** in vertexSet
 - ☐ the samllest value in weight [v]
 - You add v to vertexSet
- For all unselected vertices u not in vertexSet
 - check value of weight [u] (cost of path from 0 to u) to ensure that they are minimums
 - can the value of weight [u] be reduced by passing newly selected vertex v?

- can the value of weight [u] be reduced by passing newly selected vertex v?
- □ how to ensure value of weight [u] is minimums?
 - break the path from 0 to u into two pieces and find their weights as follows
 - □ weight [v] = weight of the shortest path from 0 to v
 - □ matrix[v][u] = weight of the edge from v to u
- \square weight[u] = min{weight[u], weight[v] + matrix[v][u]}



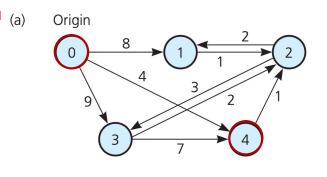


(b)					3		
	0	∞	8	∞	9	4	
	1	∞	∞	1	∞	∞	
	2	∞	2	∞	3	∞	
	3	∞	∞	2	∞	7	
	4	8 8 8	∞	1	∞	∞	

					weight			
Step	\underline{v}	<u>vertexSet</u>	[0]	[1]	[2]	[3]	[4]	
1	_	0	0	8	00	9	4	

initially, vertexSet contains 0,
weight = adjacency matrix[0][]

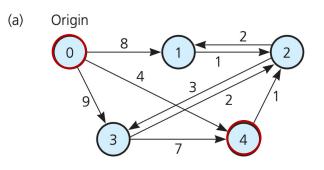
(h)



(D)		0	1	2	3	4
	0	8 8 8 8	8	∞	9	4
	1	∞	∞	1	∞	∞
	2	∞	2	∞	3	∞
	3	∞	∞	2	∞	7
	4	∞	∞	1	∞	∞

					weight			
Step	\underline{v}	vertexSet	[0]	[1]	[2]	[3]	[4]	
1	_	0	0	8	∞	9	4	
2	4	0.4						

weight[4] is the smallest value
in weight and vertex 4 didn't in
vertexSet. So, v = 4 & add 4
to vertexSet

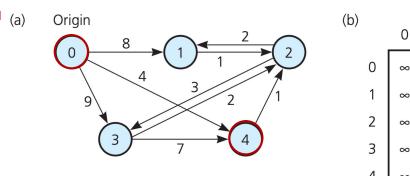


(b)		0	1	2	3	4
	0	∞	8	∞	9	4
	1	∞	∞	1	∞	∞
	2	∞	2	∞	3	∞
	3	∞	∞	2	∞	7
	4	∞	∞	∞ 1 ∞ 2 1	∞	∞

					weight	
Step	<u>v</u>	<u>vertexSet</u>	[0]	[1]	[2]	[3]
1	_	0	0	8	∞	9
2	4	0, 4	0			

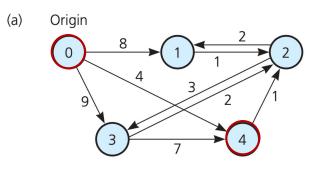
For vertices not in vertexSet (for u = 1, 2, 3), check if weight [u] (path from 0 to u) is minimums

```
weight[u] = min{weight[u],
weight[v] + matrix[v][u]}
```



					weight			
Step	\underline{v}	<u>vertexSet</u>	[0]	[1]	[2]	[3]	_[4]	For vertex 1 (u=1)
1	-	0	0	8	∞	9	4	<pre>weight[1] = 8 weight[4] + matr</pre>
2	4	0, 4	0	8			4	

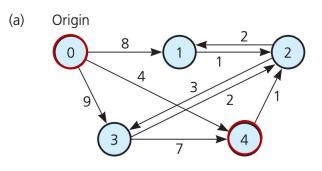
...a + a b +



0)	0	1	2	3	4
0	∞	8	∞	9	4
1	∞	∞	1	∞	∞
2	∞	2	∞	3	∞
3	∞	∞	2	∞	7
4	& & & & & & & & & & & & & & & & & & &	∞	1	∞	∞

					weight			
Step	\underline{v}	<u>vertexSet</u>	[0]	[1]	_ [2]	[3]	[4]	
1	_	0	0	8	∞	9	4	
2	4	0, 4	0	8	5		4	

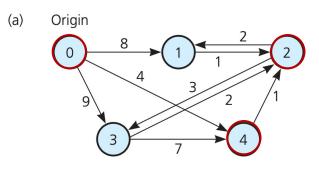
For vertex 2 (u=2), weight[2] = ∞ weight[4] + matrix[4][2] = 4 + 1



(b)		0	1	2	3	4
	0	∞	8	∞	9	4
	1	∞	∞	1	∞	∞
	2	∞	2	∞	3	∞
	3	∞	∞	2	∞	7
	4	∞	∞	1	9 ∞ 3 ∞	∞

			weight				
Step	\underline{v}	vertexSet	[0]	[1]	[2]	[3]	
1	_	0	0	8	∞	9	
2	4	0, 4	0	8	5	9	

For vertex 3 (u=3), weight[3] = 9 weight[4] + matrix[4][3] = 4 + ∞



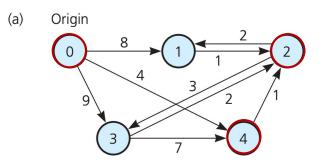
(b)		0	1	2	3	4
	0	8 8 8 8	8	∞	9	4
	1	∞	∞	1	∞	∞
	2	∞	2	∞	3	∞
	3	∞	∞	2	∞	7
	4	∞	∞	1	∞	∞
		•				

			weight						
Step	<u>v</u>	<u>vertexSet</u>	[0]	[1]	[2]	[3]	[4]		
1	-	0	0	8	∞	9	4		
2	4	0, 4	0	8	5	9	4		
3	2	-0,-4,-2	0		5		4		

weight[2] is the smallest value in
weight and vertex 2 didn't in
vertexSet.

So, v = 2 & add 2 to vertexSet

(h)

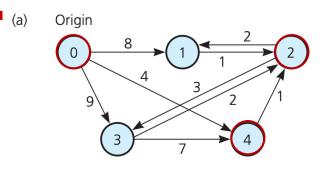


(D)		0	1	2	3	4
	0	8 8 8 8	8	∞	9	4
	1	∞	∞	1	∞	∞
	2	∞	2	∞	3	∞
	3	∞	∞	2	∞	7
	4	∞	∞	1	∞	∞

			weight					
Step	\underline{v}	vertexSet	[0]	[1]	[2]	[3]	[4]	
1	-	0	0	8	∞	9	4	
2	4	0, 4	0	8	5	9	4	
3	2	0, 4, 2	0		5		4	

For vertices not in vertexSet (for u = 1, 3), check if weight [u] (path from 0 to u) is minimums

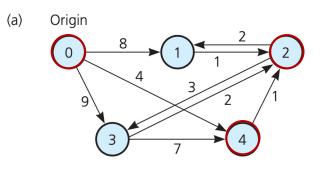
weight[u] = min{weight[u],
weight[v] + matrix[v][u]}



(b)		0			3	
	0	8	8 ∞ 2 ∞ ∞	∞	9	4
	1	∞	∞	1	∞	∞
	2	∞	2	∞	3	∞
	3	∞	∞	2	∞	7
	4	∞	∞	1	∞	∞
		•				

			weight					
Step	\underline{v}	<u>vertexSet</u>	[0]	[1]	[2]	[3]	_[4]	
1	-	0	0	8	∞	9	4	
2	4	0, 4	0	8	5	9	4	
3	2	0, 4, 2	0	7	5		4	

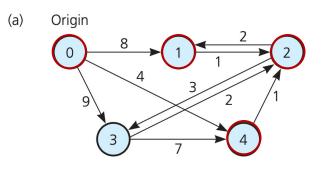
For vertex 1 (u=1), weight[1] = 8 weight[2] + matrix[2][1] = 5 + 2



(b)		0	1	2	3	4
	0	8	8	∞	9 ∞ 3 ∞	4
	1	∞	∞	1	∞	∞
	2	∞	2	∞	3	∞
	3	∞	∞	2	∞	7
	4	∞	∞	1	∞	∞
		•				

			weight					
Step	<u>v</u>	<u>vertexSet</u>	[0]	[1]	[2]	[3]	[4]	
1	-	0	0	8	∞	9	4	
2	4	0, 4	0	8	5	9	4	
3	2	0, 4, 2	0	7	5	8	4	

For vertex 3 (u=3), weight[3] = 9 weight[2] + matrix[2][3] = 5 + 3

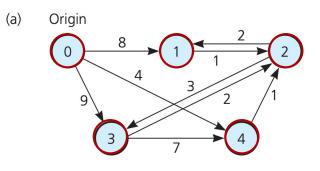


(b)		0	1	2	3	4
	0	8 8 8 8	8	∞	9	4
	1	∞	∞	1	∞	∞
	2	∞	2	∞	3	∞
	3	∞	∞	2	∞	7
	4	∞	∞	1	∞	∞

			weight					
Step	\underline{v}	<u>vertexSet</u>	[0]	[1]	[2]	[3]	[4]	
1	_	0	0	8	∞	9	4	
2	4	0, 4	0	8	5	9	4	
3	2	0, 4, 2	0	7	5	8	4	
4	1	0,-4, 2, 1	0	7	5	8	4	

weight[1] is the smallest value in weight and vertex 1 didn't in vertexSet.

So, v = 1 & add 1 to vertexSet



(b)		0	1	2	3	4
	0	8 8 8 8	8	∞	9	4
	1	∞	∞	1	∞	∞
	2	∞	2	∞	3	∞
	3	∞	∞	2	∞	7
	4	∞	∞	1	∞	∞
		-				

	weight							
	_[4]	[3]	[2]	[1]	[0]	vertexSet	\underline{v}	<u>Step</u>
repeat the same process, update the weight [u] (u = 3) the final values in weight are the weight of shortest path from vertex 0 to other vertices	4	9	∞	8	0	0	_	1
	4	9	5	8	0	0, 4	4	2
	4	8	5	7	0	0, 4, 2	2	3
	4	8	5	7	0	0, 4, 2, 1	1	4
68	4	8	5	7	0	0, 4, 2, 1, 3	3	5

```
// Finds the minimum-cost paths between an origin vertex
// (vertex 0) and all other vertices in a weighted directed
shortestPath(theGraph: Graph, weight: WeightArray)
     // Step 1: initialization
     Create a set vertexSet that contains only vertex 0
     n = number of vertices in theGraph
     for (v = 0 \text{ through } n - 1)
           weight[v] = matrix[0][v]
     // Steps 2:
     for (step = 2 through n) { //for the rest vertex 1 to n-1
           Find the smallest weight[v] such that v is not in vertexSet
           Add v to vertexSet
           for (all vertices u not in vertexSet) {// update weight[u] for all u not in vertexSet
                if (weight[u] > weight[v] + matrix[v][u])
                      weight[u] = weight[v] + matrix[v][u]
```

Summary of Graph

- Graph implementation
 - adjacency matrix and the adjacency list
 - advantages and disadvantages
- Graph Traversal
 - Depth-first search (DFS)
 - Breadth-first search (BFS)
 - the order of visit
- Applications
 - Topological sorting
 - Dijkstra's shortest-path algorithm