CSE 100
Graphs
Kinds of Data Structures

Unstructured structures (sets)

Sequential, linear structures (arrays, linked lists)

Hierarchical structures (trees)

Graphs

Consist of:
- A collection of elements (“nodes” or “vertices”)
- A set of connections (“edges” or “links” or “arcs”) between pairs of nodes.
  - Edges may be directed or undirected
  - Edges may have weight associated with them

Graphs are not hierarchical or sequential, no requirements for a “root” or “parent/child” relationships between nodes
Kinds of Data Structures

Unstructured structures (sets)
Sequential, linear structures (arrays, linked lists)
Hierarchical structures (trees)

Graphs
A. They consist of both vertices and edges
B. They do NOT have an inherent order
C. Edges may be weighed or unweighted
D. Edges may be directed or undirected
E. They may contain cycles
Kinds of Data Structures

Unstructured structures (sets)
Sequential, linear structures (arrays, linked lists)
Hierarchical structures (trees)

Graphs
Which of the following is true?
A. A graph can always be represented as a tree
B. A tree can always be represented as a graph
C. Both A and B
D. Neither A or B
Kinds of Data Structures

**Unstructured structures** (sets)

**Sequential, linear structures** (arrays, linked lists)

**Hierarchical structures** (trees)

**Graphs**
Which of the following is true?
A. A graph can always be represented as a tree
B. A tree can always be represented as a graph
C. Both A and B
D. Neither A or B

Note that trees are special cases of graphs; lists are special cases of trees.
Why Graphs?

Many real-world systems map well to graphs
Why Graphs?

Remember: If your problem maps to a well-known graph problem, it usually means you can solve it blazingly fast!

$O(|V| + |E|)$

This is considered fast
A directed graph

\[ \text{directed } G = (V, E) \]

\[ V = \{ v_i : i = 0 \text{ to } 5 \} \]

\[ N = |V| = 6 \]

\[ E = \{ (v_0, v_1), (v_1, v_4), (v_4, v_1), (v_1, v_3), (v_3, v_2), (v_2, v_5), (v_3, v_4) \} \]

\[ m = |E| = 8 \]

Path: \( \{ v_2, v_0, v_1, v_4 \} \)
A graph $G = (V,E)$ consists of a set of vertices $V$ and a set of edges $E$

- Each edge in $E$ is a pair $(v,w)$ such that $v$ and $w$ are in $V$.
- If $G$ is an *undirected* graph, $(v,w)$ in $E$ means vertices $v$ and $w$ are connected by an edge in $G$. This $(v,w)$ is an unordered pair.
- If $G$ is a *directed* graph, $(v,w)$ in $E$ means there is an edge going from vertex $v$ to vertex $w$ in $G$. This $(v,w)$ is an ordered pair; there may or may not also be an edge $(w,v)$ in $E$.
- In a *weighted* graph, each edge also has a “weight” or “cost” $c$, and an edge in $E$ is a triple $(v,w,c)$.
- When talking about the size of a problem involving a graph, the number of vertices $|V|$ and the number of edges $|E|$ will be relevant.
Connected, disconnected and fully connected graphs

- Connected graphs:
  - (undirected) path between every pair of vertices

- Disconnected graphs:
  - Weakly connected: If the undirected version of the graph is connected
  - Strongly connected: If a path exists between every vertex pair
  - A graph that is not connected is disconnected

- Fully connected (complete graphs):
  - There exists an edge between every vertex pair
Q: What are the minimum and maximum number of edges in a undirected connected graph $G(V,E)$ with no self loops, where $N=|V|$?

A. 0, $N^2$
B. $N$, $N^2$
C. $N-1$, $N(N-1)/2$
A dense graph is one where $|E|$ is “close to” $|V|^2$. A sparse graph is one where $|E|$ is “closer to” $|V|$.
Representing Graphs: Adjacency Matrix

A 2D array where each entry $[i][j]$ encodes connectivity information between $i$ and $j$

- For an unweighted graph, the entry is 1 if there is an edge from $i$ to $j$, 0 otherwise
- For a weighted graph, the entry is the weight of the edge from $i$ to $j$, or “infinity” if there is no edge
- Note an undirected graph’s adjacency matrix will be symmetrical
How big is an adjacency matrix in terms of the number of nodes and edges (BigO, tightest bound)?

A. $|V|$
B. $|V| + |E|$
C. $|V|^2$
D. $|E|^2$
E. Other

When is that OK? When is it a problem?
Space efficiency of Adjacency Matrix

A dense graph is one where $|E|$ is “close to” $|V|^2$.
A sparse graph is one where $|E|$ is “closer to” $|V|$.

Adjacency matrices are space inefficient for sparse graphs.
Representing Graphs: Adjacency Lists

- Vertices and edges stored as lists
- Each vertex points to all its edges
- Each edge points to the two vertices that it connects
- If the graph is directed: edge nodes differentiate between the head and tail of the connection
- If the graph is weighted edge nodes also contain weights

Vertex List

Edge List
Representing Graphs: Adjacency Lists

Each vertex has a list with the vertices adjacent to it. In a weighted graph this list will include weights.

How much storage does this representation need? (BigO, tightest bound)

A. $|V|$
B. $|E|$
C. $|V| + |E|$
D. $|V|^2$
E. $|E|^2$

For sparse graphs the adjacency list representation is more space efficient.
Searching a graph

- Find if a path exists between any two nodes
- Find the shortest path between any two nodes
- Find all nodes reachable from a given node

Generic Goals:
- Find everything that can be explored
- Don’t explore anything twice
Generic approach to graph search
Depth First Search for Graph Traversal

- Search as far down a single path as possible before backtracking
Depth First Search for Graph Traversal

• Search as far down a single path as possible before backtracking

Assuming DFS chooses the lower number node to explore first, in what order does DFS visit the nodes in this graph?

A. V0, V1, V2, V3, V4, V5
B. V0, V1, V3, V4, V2, V5
C. V0, V1, V3, V2, V4, V5
D. V0, V1, V2, V4, V5, V3
Depth First Search for Graph Traversal

- Search as far down a single path as possible before backtracking

Does DFS always find the shortest path between nodes?
A. Yes
B. No