Graph Theory

Nirdosh Bhatnagar

Disclaimer: *These notes are neither thorough nor complete. They may be distributed outside the class only with the permission of the Instructor.*

1. Definitions

A graph is generally used to model the connectivity of a system. A graph is made up of dots, called vertices and the lines connecting them. These lines are called edges or links. If these lines are not assigned a direction, then the corresponding graph is called an **undirected graph**. However, if these lines are assigned a direction, the associated graph is called a **directed graph** or simply a **digraph**.

1.1. Undirected graph.

An **undirected graph** G is made up of two sets V and E. The set V consists of **vertices**, and the set E consists of **edges**. A member of the set V is a vertex. It is also referred to as a **node**. A vertex or a node is typically represented as a point, while an edge is represented as a line. An edge joins either a pair of vertices or joins a vertex to itself. This graph is denoted as $G = (V, E)$.

A **proper edge** joins one vertex to another, while a **self-loop** or simply a **loop** joins a vertex to itself. A proper edge has two different **endpoints**, while a loop has a single endpoint.

The cardinality of sets V and E are denoted by $|V|$ and $|E|$ respectively. $|V|$ is called the **order** of G, and $|E|$ is called the **size** of G.

An edge e is defined by (u, v) where $u, v \in V$. Observe that no direction has been assigned to the edge, that is the edge (u, v) of an undirected graph can also be represented by (v, u).

Vertices u and v are **adjacent** to each other if there is an edge $e = (u, v)$. The edge e is said to be **incident** upon vertices u and v. Adjacent vertices are also called **neighbors**.

The **degree of a vertex** v is the number of edges incident upon v. An **isolated point** of a graph G is a vertex which is not joined by an edge to any other vertex of the graph.

A **pendant vertex** or a **leaf** is a vertex with degree one.

A **weighted graph** has a number assigned to each edge. This number is called the weight of the edge.

A **walk** in the graph G is a sequence of nodes $u_1, u_2, \ldots, u_m$ where $m \geq 1$ and $(u_j, u_{j+1}) \in E$ for $1 \leq j \leq (m-1)$. The **length of the walk** is $(m-1)$. The walk is considered to be **closed** if $u_1 = u_m$ and m is greater than 1. If the walk has no repeated nodes, then it is called a **path**. A **circuit** or a **cycle** is $u_1, u_2, \ldots, u_m = u_1$ where $m > 1$ and no repeated intermediate nodes.

A **multi-edge** in an undirected graph G is a set of more than one edge between the same pair of vertices. If an undirected graph has multiple edges, then the graph is called a **multi-graph**.

A graph without multiple edges and loops is called a **simple graph**.

1.2. Directed graph.

Directions are assigned to an edge in a directed graph. In this case the corresponding graph is called a directed graph, or a digraph. A digraph H is a pair $H = (V, E)$ where V is a set of vertices and E is a set of ordered pairs of vertices called **arcs**. An arc is simply a directed edge.

The **tail of an arc** is the vertex at which the arc originates and the **head of an arc** is the vertex at which the arc terminates. Let an arc be $e = (u, v)$ where $u, v \in V$. Then the vertices u and v are the tail and head of the arc e respectively.

The **out-degree of a vertex** v is the number of arcs belonging to the set E that are (v, u). Similarly, the **in-degree of a vertex** v is the number of arcs belonging to the set E that are (u, v).

Denote the out-degree of a vertex v, by $d^+(v)$. It is equal to the number of arcs with tail at v. Also denote the in-degree of a vertex v, by $d^-(v)$. It is equal to the number of arcs with head at v.

A **directed walk** in the graph H is a sequence of nodes $u_1, u_2, \ldots, u_m$ where $m \geq 1$ and $(u_j, u_{j+1}) \in E$ for $1 \leq j \leq (m-1)$. The **length of the walk** is $(m-1)$. The directed walk is considered to be **closed** if $u_1 = u_m$ and m is greater than 1. If this walk has no repeated nodes, then it is called a **directed path**. A **directed circuit or a cycle** is $u_1, u_2, \ldots, u_m = u_1$ where $m > 1$ and no repeated intermediate nodes.

A multi-edge in a directed graph G is a set of more than one arc between the same pair of vertices. These arcs have the same head and tails. The corresponding graph is called **directed multi-graph**.

A directed graph without multiple edges and loops is called a **strict digraph**.

2. Special graphs

In a **null graph**, both the vertex set and the edge set are empty.

K_n is a **complete graph** with n vertices, where all pairs of vertices are adjacent. A complete graph is also called a **clique**.

A graph is **connected** if every pair of nodes is connected by a path.

A graph is **disconnected**, if there is at least one pair of nodes which cannot be linked. A disconnected graph can be split into several components in which each **component** graph is connected.

The **diameter** of a connected graph is defined as the largest distance between any two vertices of the graph. In this context the distance between any two nodes is assumed to be the number of edges in the shortest path between the two vertices.

A graph is said to be **labeled**, if each vertex in the graph is assigned a unique label.

A **tree** is a connected undirected graph without cycles. Let $G = (V, T)$ be a tree graph, then G is said to **span** its set of nodes V.

A tree in which one vertex is specifically identified as a **root-node** is called the **rooted-tree**.

If G is any undirected graph, then a **spanning tree** is a tree that spans all the nodes in the vertex set V.

A **forest** is graph with an ensemble of trees.

3. Graph operations

A **subgraph** of a graph $G = (V, E)$ is a graph $G_s = (V_s, E_s)$ where $V_s \subseteq V$ and $E_s \subseteq E$. Note that if $e = (u, v) \in E_s$ then $u, v \in V_s$.

A **spanning subgraph** of G is a subgraph where $V = V_s$.

Deletion of an edge e in a graph G results in a subgraph denoted by $G - e$. This subgraph contains all the vertices of graph G, and all edges of G except the edge e.

Deletion of a vertex v in a graph G results in a subgraph denoted by $G - v$. This subgraph contains all the vertices of graph G except v, and all the edges of G except those incident upon the vertex v.

Contraction of an edge e in a graph G results in a graph in which the edge e is shrunk to a point. That is the endpoints of the edge are merged to form a single vertex. The rest of the graph remains unchanged. This graph is denoted by G/e.

4. Graph Representation

We outline different representations of a graph.

1. **Endpoint table:** This is a representation of the graph in a table. It gives the endpoints of all edges.

   We illustrate this by an example. Consider an undirected graph $G = (V, E)$, where $V = \{v_1, v_2, v_3, v_4\}$, $E = \{e_1, e_2, e_3, e_4, e_5, e_6\}$, and $e_1 = (v_1, v_3)$, $e_2 = (v_2, v_4)$, $e_3 = (v_2, v_3)$, $e_4 = (v_1, v_4)$, $e_5 = (v_1, v_4)$, and $e_6 = (v_1, v_1)$. The endpoint table of this digraph is

e_1	e_2	e_3	e_4	e_5	e_6
v_1	v_2	v_2	v_1	v_1	v_1
v_3	v_4	v_3	v_4	v_4	

   The representation is similar for a digraph.

2. **Incident-edge table:** This is a tabular representation of the graph, where for each vertex v, a list of the edges incident upon this vertex is given. For a digraph, there are two sublists of edges for each vertex. The lists correspond to whether v is a tail or a head.

   The incident-edge table of the above graph is

   $$\begin{array}{lllll} v_1: & e_1 & e_4 & e_5 & e_6 \\ v_2: & e_2 & e_3 & & \\ v_3: & e_1 & e_3 & & \\ v_4: & e_2 & e_4 & e_5 & \end{array}$$

3. **List of neighbors representation:** For each vertex v, the immediate neighbors of the vertex v are listed.

   The list of neighbors representation of the above graph is

   $$\begin{array}{lllll} v_1: & v_1 & v_3 & v_4 & v_4 \\ v_2: & v_3 & v_4 & & \\ v_3: & v_1 & v_2 & & \\ v_4: & v_1 & v_1 & v_2 & \end{array}$$

4. **Adjacency set**: The adjacency set A_v, of a vertex v is the set of all arcs leaving this vertex. $A_v = \{(v, v_j) \mid (v, v_j) \in E\}$.

5. **Adjacency matrix:** Let G be an undirected graph (V, E). Let $V = \{v_1, v_2, \ldots, v_n\}$ and $E = \{e_1, e_2, \ldots, e_m\}$. The adjacency matrix A is an $n \times n$ matrix. Let the element of the matrix be $a_{ij}, 1 \leq i, j \leq n$. Then

   $$a_{ij} = \begin{cases} \text{number of edges between } v_i \text{ and } v_j & \text{if } i \neq j \\ \text{number of self-loops} & \text{if } i = j \end{cases}$$

If the graph is directed, that is $H = (V, E)$, then the elements of the corresponding adjacency matrix A are

$$a_{ij} = \text{ number of arcs from } v_i \text{ to } v_j$$

A row-sum in the adjacency matrix of a digraph is equal to the out-degree of the corresponding vertex. And the column-sum is equal to the in-degree. For the above graph the adjacency matrix is

$$A = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \begin{array}{c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

6. **Incidence matrix:** Let G be an undirected graph (V, E). Let $V = \{v_1, v_2, \ldots, v_n\}$ and $E = \{e_1, e_2, \ldots, e_m\}$. The incidence matrix I_G is an $n \times m$ matrix. Let the element of the matrix be $b_{ij}, 1 \leq i \leq n$ and $1 \leq j \leq m$. Then

$$b_{ij} = \begin{cases} 0, & \text{if } v_i \text{ is not an endpoint of } e_j \\ 1, & \text{if } v_i \text{ is an endpoint of a proper edge } e_j \\ 2, & \text{if } e_j \text{ is a self-loop at } v_i \end{cases}$$

In this matrix every row-sum is equal to the degree of the corresponding vertex, and the column-sum is equal to 2. If the graph is directed, that is $H = (V, E)$, then the elements of the corresponding incidence matrix I_H are

$$b_{ij} = \begin{cases} +1, & \text{if } v_i \text{ is the tail of } e_j \\ -1, & \text{if } v_i \text{ is the head of } e_j \\ 0, & \text{otherwise} \end{cases}$$

Occurrence of self-loops is not defined. The incidence matrix of the graph in the example is

$$I_G = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \begin{array}{c} \begin{array}{cccccc} e_1 & e_2 & e_3 & e_4 & e_5 & e_6 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

Observe that the row-sums are equal to the degree of the corresponding vertex, and all the column-sums are equal to 2.

5. Observations

We list some observations about both undirected and directed graphs.

5.1. Undirected graph observations.

1. **Euler's theorem:** In a graph G, the sum of the degrees of vertices is equal to twice the number of edges.
2. In a graph G, the number of vertices of odd degree is even.
3. A tree with the number of vertices, $n \geq 2$ has at least two leaves.
4. In computer systems, a graph is usually represented by an endpoint table and the incident-edge table.
5. **Matrix-tree theorem:** Let G be a graph with an adjacency matrix A. Also D is a diagonal matrix, with the degrees of vertices on its diagonal. Then the total number of spanning trees of the graph is equal to any cofactor of the matrix $(D - A)$.
6. Let A be the adjacency matrix of a simple graph G with n vertices. Then the ijth element of the matrix $A^k, 1 \leq k \leq (n-1)$ is the number of different edge sequence of k edges between the vertices v_i and v_j.
7. Let A be the adjacency matrix of a graph G with n vertices. Define
$$Y = \sum_{k=1}^{(n-1)} A^k$$
Then the graph G is a disconnected graph if and only if there exists at least one element of the matrix Y that is zero. Use of this equation to test the connectedness of a graph is not computationally efficient.
8. A graph is said to be disconnected if and only if its vertices can be ordered such that its adjacency matrix A can be represented as
$$A = \begin{bmatrix} A_1 & \varnothing \\ \varnothing & A_2 \end{bmatrix}$$
where A_1 and A_2 are adjacency matrices of disconnected subgraphs G_1 and G_2 respectively. That is there is no edge joining any pair of vertices in the two subgraphs.
9. Some properties of a tree are summarized below
   (a) The number of edges in a tree with n vertices is equal to $(n-1)$.
   (b) The tree has no cycles.
   (c) There is no more than a single path between a pair of vertices of a tree.
   (d) The total number of labeled trees with n vertices is n^{n-2}, for $n \geq 2$

5.2. Digraph observations.

1. A digraph can be specified by an arc list, an incidence matrix, adjacency matrix, or list of neighbors in a computer system.
2. In any digraph $H = (V, E)$, the sum of in-degrees, and the sum of out-degrees are each equal to $|E|$, the number of arcs.
$$\sum_{v \in V} d^-(v) = \sum_{v \in V} d^+(v) = |E|$$