Fundamentals of Data Structure

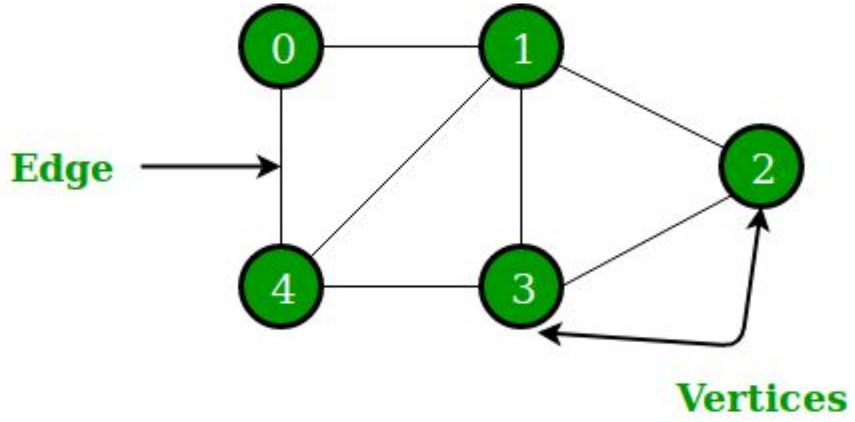
Mahesh Shirole

VJTI, Mumbai-19

Slides are prepared from

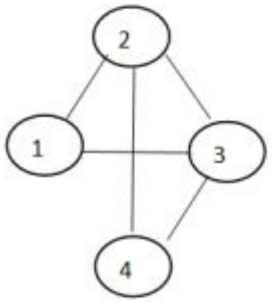
1. Data Structures and Algorithms in Java, 6th edition, by M. T. Goodrich, R. Tamassia, and M. H. Goldwasser, Wiley, 2014
2. Data Structures and Algorithms in Java, by Robert Lafore, Second Edition, Sams Publishing

Graph

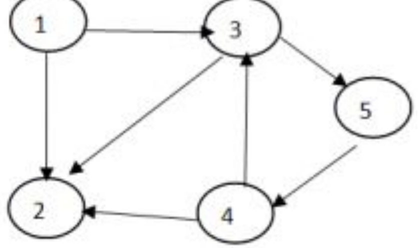


- Graphs are one of the most versatile structures used in computer programming
- The sorts of problems that graphs can help to solve are generally quite different
- A graph is a way of representing relationships that exist between pairs of objects
- A graph is a set of objects, called ***vertices***, together with a collection of pairwise connections between them, called ***edges***
- The circles are vertices, and the lines are edges. The vertices are usually labeled with alphabets or numbers.
- Graphs have applications in modeling many domains, including mapping, transportation, computer networks, and electrical engineering

Graph



Undirected graph

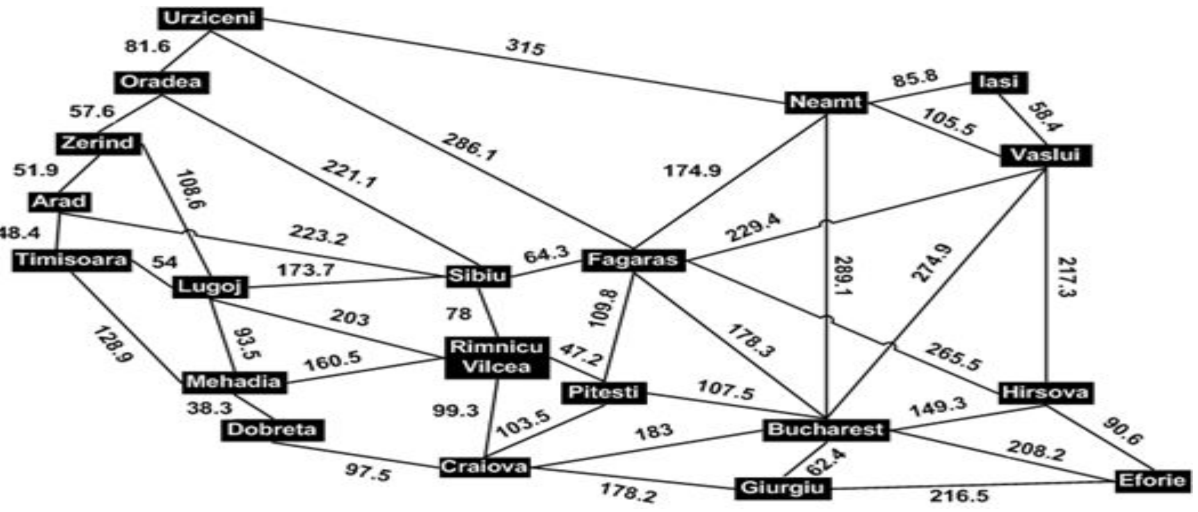


Directed graph.

- A graph ***G (V, E)*** is simply a set ***V*** of vertices and a collection ***E*** of pairs of vertices from ***V***, called edges
- Thus, a graph is a way of representing connections or relationships between pairs of objects from some set ***V***
- Edges in a graph are either *directed* or *undirected*
- An edge ***(u,v)*** is said to be directed from ***u*** to ***v*** if the pair ***(u,v)*** is ordered, with ***u*** preceding ***v***.
- An edge ***(u,v)*** is said to be undirected if the pair ***(u,v)*** is not ordered
- If all the edges in a graph are *undirected*, then we say the graph is an ***undirected graph***
- If all the edges in a graph are *directed*, then we say the graph is an ***directed graph***, also called a ***digraph***
- A graph that has both directed and undirected edges is often called a ***mixed graph***

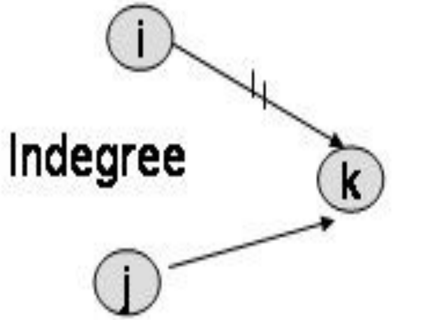
Graph Examples

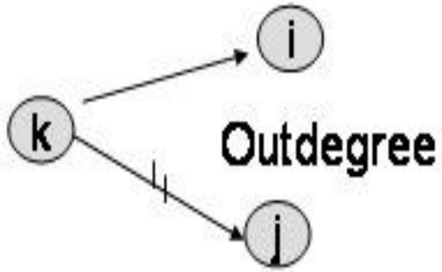
- **A city map** can be modeled as a graph whose vertices are intersections or dead ends, and whose edges are stretches of streets without intersections. This graph has both undirected edges, which correspond to stretches of two-way streets, anddirectededges, whichcorrespondtostretches ofone-waystreets. Thus, in this way, a graph modeling a city map is a mixed graph

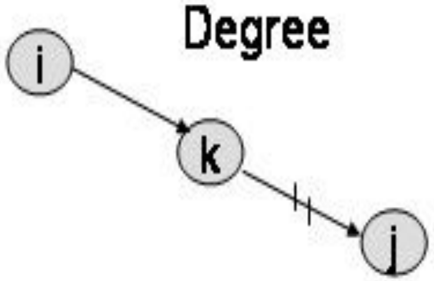


Graph

- The two vertices joined by an edge are called the end vertices (or ***endpoints***) of the edge
- If an edge is directed, its first end point is its ***origin*** and the other is the ***destination*** of the edge
- Two vertices ***u*** and ***v*** are said to be **adjacent** if there is an edge whose end vertices are ***u*** and ***v***
- An edge is said to be **incident** to a vertex if the vertex is one of the edge's endpoints
- The outgoing edges of a vertex are the directed edges whose origin is that vertex
- The incoming edges of a vertex are the directed edges whose destination is that vertex
- The degree of a vertex ***v***, denoted ***deg(v)***, is the number of incident edges of ***v***
- The in-degree and out-degree of a vertex ***v*** are the number of the incoming and outgoing edges of ***v***, and are denoted ***indeg(v)*** and ***outdeg(v)***, respectively

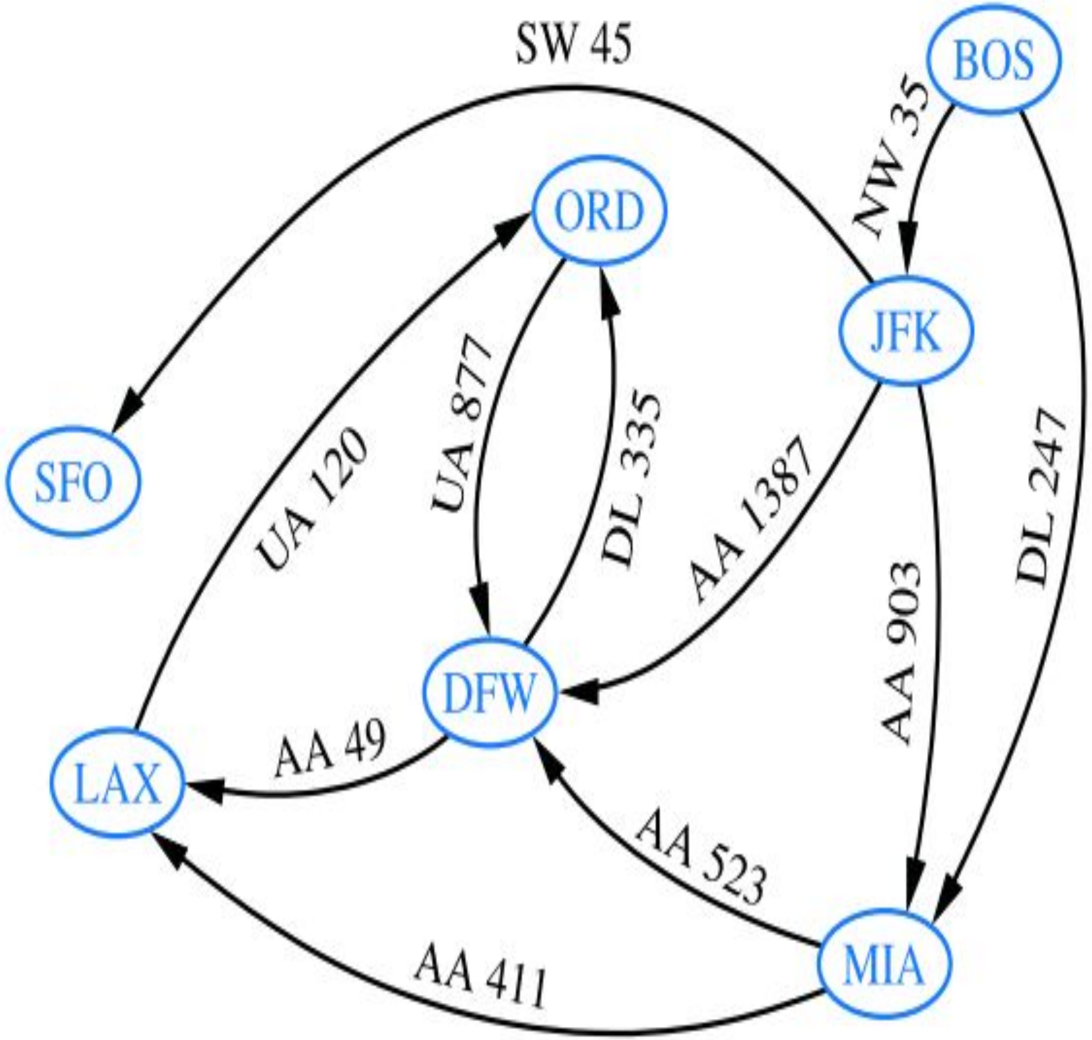




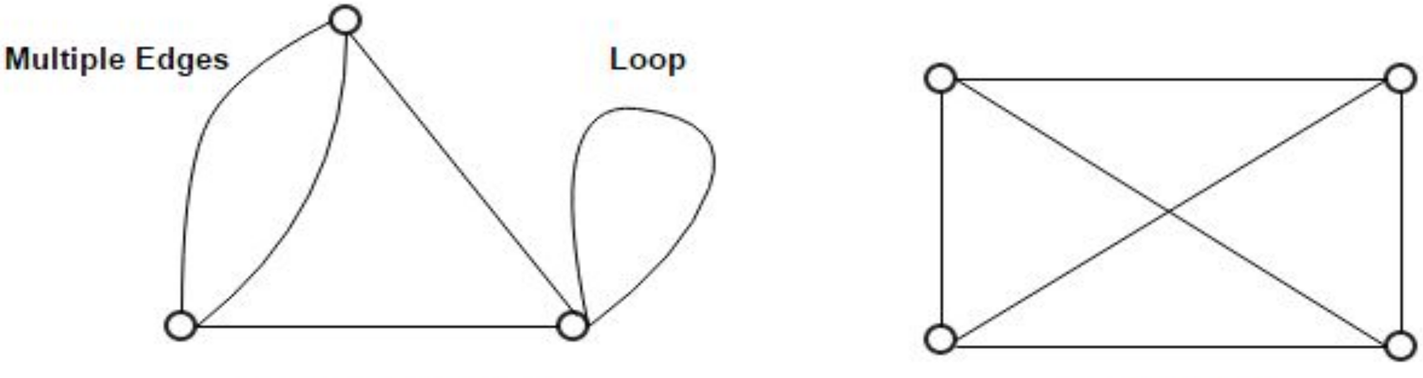


Graph Example

- We can study air transportation by constructing a graph ***G***, called a flight network, whose vertices are associated with airports, and whose edges are associated with flights
- Two airports are adjacent in ***G*** if there is a flight that flies between them, and an edge ***e*** is incident to a vertex ***v*** in ***G*** if the flight for ***e*** flies to or from the airport for ***v***
- The in-degree of a vertex ***v*** of ***G*** corresponds to the number of inbound flights to ***v***'s airport, and the out-degree ofa vertex ***v*** in ***G*** corresponds to the number of out bound flights



Graph

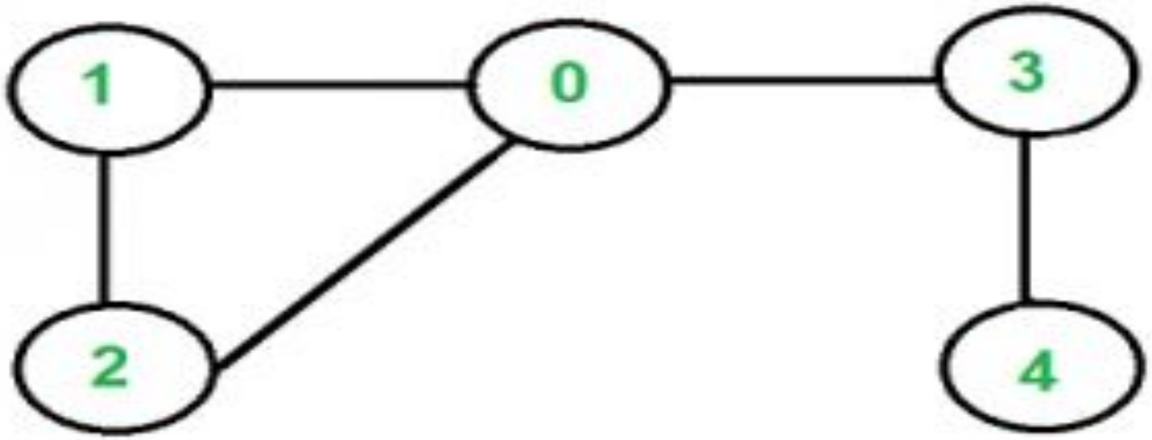


Not a Simple Graph

Simple Graph

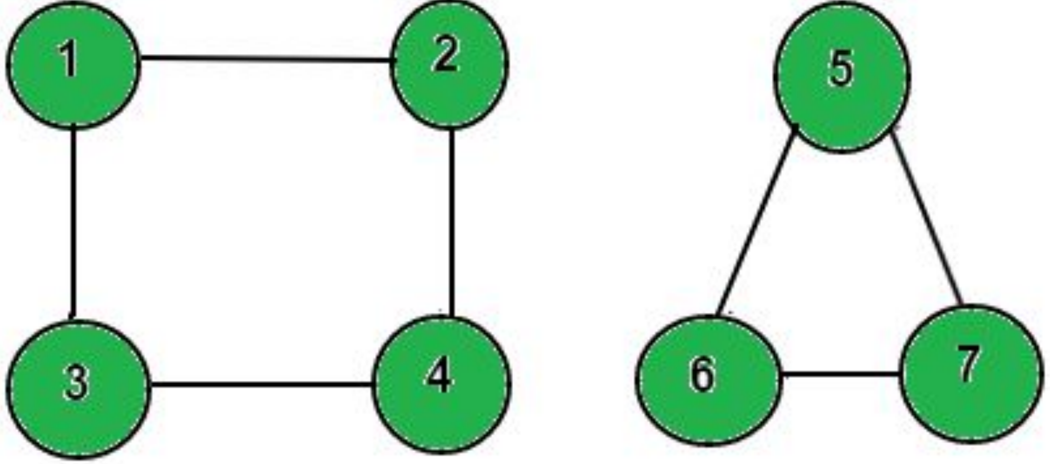
- The definition of a graph refers to the group of edges as a collection, not a set, thus allowing two undirected edges to have the same end vertices, and for two directed edges to have the same origin and the same destination
- Such edges are called **parallel edges** or **multiple edges**
- Another special type of edge is one that connects a vertex to itself
- Namely, we say that an edge (undirected or directed) is a **self-loop** if its two endpoints coincide
- Graphs do not have *parallel edges* or *self-loops* are said to be **simple graphs**

Graph



- A **path** is a sequence of alternating vertices and edges that starts at a vertex and ends at a vertex such that each edge is incident to its predecessor and successor vertex
- A **cycle** is a path that starts and ends at the same vertex, and that includes at least one edge
- We say that a *path* is **simple** if *each vertex in the path is distinct*
- We say that a *cycle* is **simple** if *each vertex in the cycle is distinct, except for the first and last one*
- A **directed path** is a path such that all edges are directed and are traversed along their direction
- A *directed graph* is **acyclic** if it has no directed cycles

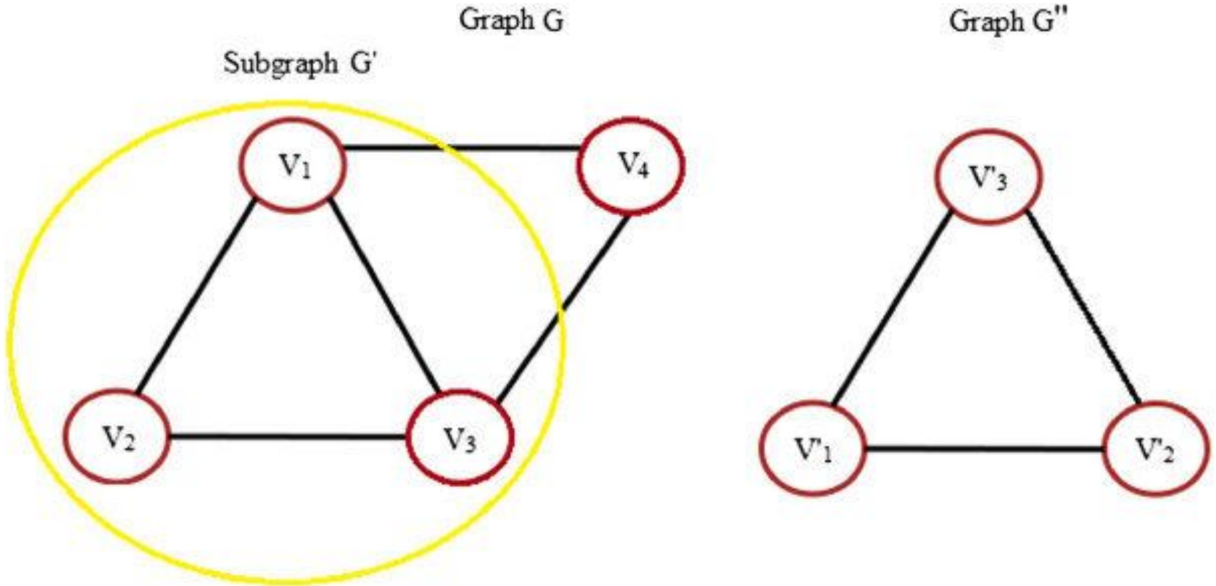
Graph

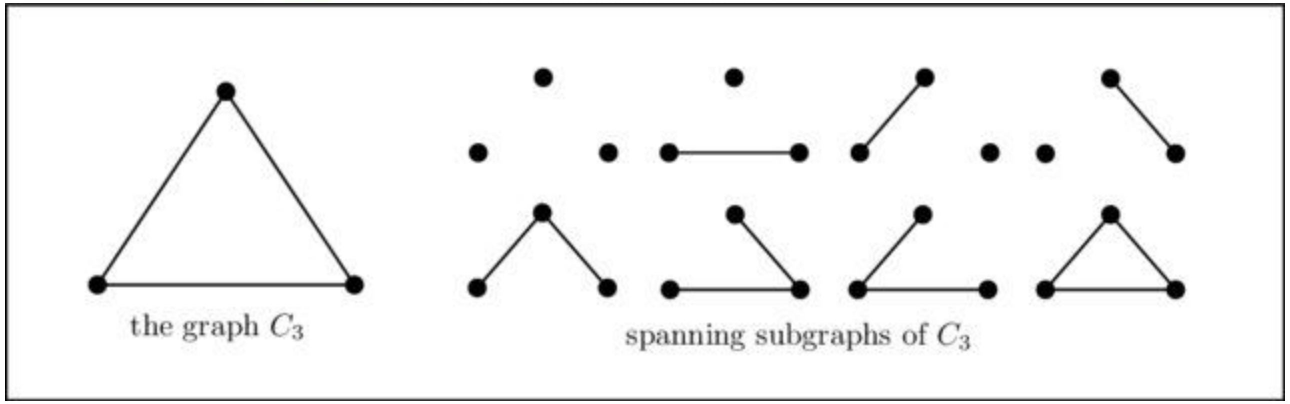


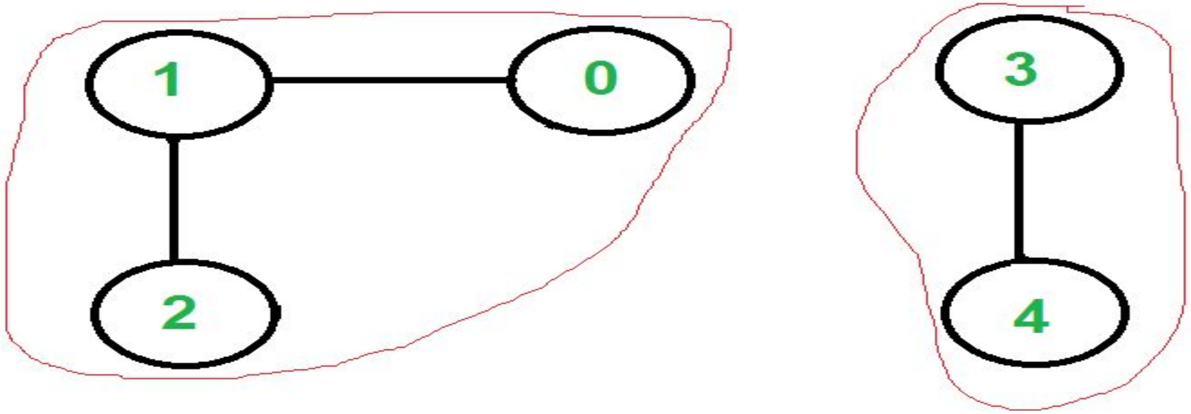
- Given vertices ***u*** and ***v*** of a (directed) graph ***G***, we say that ***u*** reaches ***v***, and that ***v*** is reachable from ***u***, if ***G*** has a (directed) **path** from ***u*** to ***v***
- In an undirected graph, the notion of reachability is symmetric, that is to say, ***u*** reaches ***v*** if an only if ***v*** reaches ***u***
- However, in a directed graph, it is possible that ***u*** reaches ***v*** but ***v*** does not reach ***u***, because a directed path must be traversed according to the respective directions of the edges
- A graph is **connected** if, for any two vertices, there is a path between them
- A directed graph G is **strongly connected** if for any two vertices ***u*** and ***v*** of ***G***, ***u*** reaches ***v*** and ***v*** reaches ***u***

Graph

- A **subgraph** of a graph ***G*** is a graph ***H*** whose vertices and edges are subsets of the vertices and edges of G, respectively
- A **spanning subgraph** of ***G*** is a subgraph of ***G*** that contains all the vertices of the graph ***G***
- If a graph ***G*** is not connected, its maximal connected subgraphs are called the connected components of ***G***
- A **forest** is a graph without cycles
- A **tree** is a connected forest, that is, a connected graph without cycles
- A **spanning tree** of a graph is a spanning subgraph that is a tree







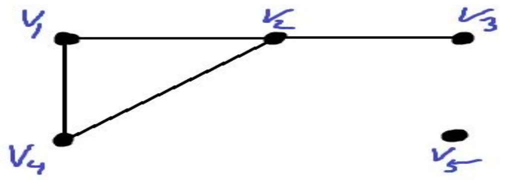
Graph

Proposition 14.8: *If G is a graph with m edges and vertex set V, then*

$$\sum_{v \text{ in } V} deg(v) = 2m.$$

Justification: An edge (u, v) is counted twice in the summation above; once by its endpoint u and once by its endpoint v. Thus, the total contribution of the edges to the degrees of the vertices is twice the number of edges. ■

Proposition 14.9: *If G is a directed graph with m edges and vertex set V, then*



$deg(v_1) = 2$
$deg(v_2) = 3$
$deg(v_4) = 2$

$deg(v_3) = 1$
$deg(v_5) = 0$

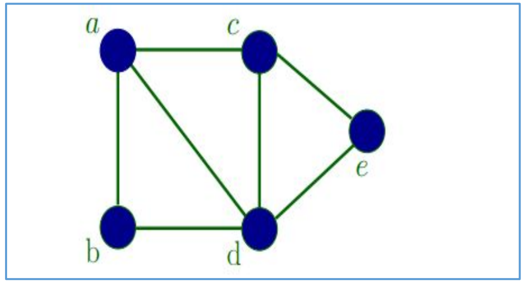
$$\sum_{v \text{ in } V} indeg(v) = \sum_{v \text{ in } V} outdeg(v) = m.$$

Justification: In a directed graph, an edge (u, v) contributes one unit to the out-degree of its origin u and one unit to the in-degree of its destination v. Thus, the total contribution of the edges to the out-degrees of the vertices is equal to the number of edges, and similarly for the in-degrees. ■

Proposition 14.10: *Let G be a simple graph with n vertices and m edges. If G is undirected, then $m \le n(n-1)/2$, and if G is directed, then $m \le n(n-1)$.*

Graph

Justification: Suppose that G is undirected. Since no two edges can have the same endpoints and there are no self-loops, the maximum degree of a vertex in G is $n-1$ in this case. Thus, by Proposition 14.8, $2m \le n(n-1)$. Now suppose that G is directed. Since no two edges can have the same origin and destination, and there are no self-loops, the maximum in-degree of a vertex in G is $n-1$ in this case. Thus, by Proposition 14.9, $m \le n(n-1)$. ■



There are a number of simple properties of trees, forests, and connected graphs.

Proposition 14.11: *Let G be an undirected graph with n vertices and m edges.*

- *If G is connected, then $m \ge n-1$.*
- *If G is a tree, then $m = n-1$.*
- *If G is a forest, then $m \le n-1$.*