Graphs for MAT 1348

by Mateja Šajna
University of Ottawa

Contents

References

[1] J. A. Bondy, U. S. R. Murty, *Graph theory with applications*, American Elsevier Publishing, New York, 1976.

[2] J. A. Bondy, U. S. R. Murty, *Graph theory.* Graduate Texts in Mathematics **244**, Springer, New York, 2008.

[3] Kenneth H. Rosen, *Discrete Mathematics and Its Applications*, McGraw-Hill, 2011.

1 Introduction

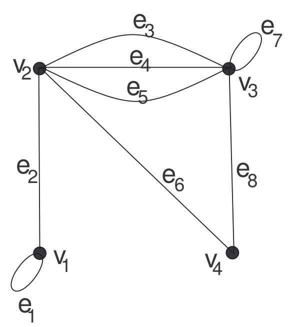
1.1 Graphs

Definition 1.1 A **graph** G is an ordered pair (V, E), where

- $V = V(G)$ is a non-empty set of **vertices** — the **vertex set** of G;
- $E = E(G)$ is a set of **edges** — the **edge set** of G; and
- the two sets are related through a function

$$\psi_G : E \to \{\{u, v\} : u, v \in V\},$$

called the **incidence function**, assigning to each edge the *unordered* pair of its **endpoints**.



Example 1.2 The drawing above represents a graph with vertex set $V = \{v_1, v_2, v_3, v_4\}$, edge set $E = \{e_1, e_2, \dots, e_8\}$, and incidence function defined by

$$\begin{aligned} \psi(e_1) &= \{v_1\} & \psi(e_5) &= \{v_2, v_3\} \\ \psi(e_2) &= \{v_1, v_2\} & \psi(e_6) &= \{v_2, v_4\} \\ \psi(e_3) &= \{v_2, v_3\} & \psi(e_7) &= \{v_3\} \\ \psi(e_4) &= \{v_2, v_3\} & \psi(e_8) &= \{v_3, v_4\} \end{aligned}$$

Definition 1.3 An edge e in a graph G is called a

- **loop** if $\psi_G(e) = \{u\}$ for some vertex $u \in V(G)$ (that is, if its endpoints coincide);
- **link** if $\psi_G(e) = \{u, v\}$ for distinct vertices $u, v \in V(G)$.

Distinct edges e_1 and e_2 in a graph G are called **parallel** or **multiple** if $\psi_G(e_1) = \psi_G(e_2)$, that is, if they have the same endpoints.

Example 1.4 In Example 1.2, edges e_1 and e_7 are loops, and all other edges are links. Edges e_4, e_5, and e_6 are pairwise parallel. Draw an edge parallel to e_7.

Definition 1.5 A **simple graph** is a graph without loops and without multiple edges.

Note: In a simple graph, the incidence function may be omitted; that is, we may identify an edge e with its unordered pair of endpoints. Thus,

$$\text{we write} \quad e = \{u, v\} \quad \text{instead of} \quad \psi(e) = \{u, v\}.$$

This may be done without ambiguity even in non-simple graphs when there is no need to distinguish between distinct parallel edges.

Moreover,

$$\text{we write shortly} \quad e = uv \quad \text{instead of} \quad e = \{u, v\},$$

omitting the braces. Note that uv is an *unordered* pair of vertices u and v; thus, $uv = vu$.

1.2 Directed Graphs (Digraphs)

Definition 1.6 A **directed graph** (or **digraph**) D is an ordered pair (V, A), where

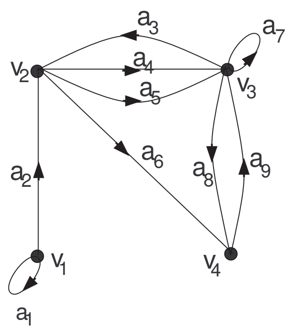
- $V = V(D)$ is a non-empty set of **vertices** — the **vertex set** of D;
- $A = A(D)$ is a set of **arcs** or **directed edges** — the **arc set** of D; and
- the two sets are related via an **incidence function** $\psi_D : A \to V \times V$, assigning to each arc the *ordered* pair of its endpoints.

If $a \in A(D)$ and $u, v \in V(D)$ are such that $\psi_D(a) = (u, v)$, then u is called the **initial** and v is called the **terminal** vertex of the arc a.

Distinct arcs a_1 and a_2 in a digraph D are called **parallel** or **multiple** if $\psi_D(a_1) = \psi_D(a_2)$, that is, if they have the same initial and the same terminal vertex.

An arc $a \in A(D)$ is called a **directed loop** if $\psi_D(a) = (u, u)$ for some vertex $u \in V(D)$.

Note: The incidence function may be omitted when a digraph has no directed loops and no multiple arcs, as well as when there is no need to distinguish between distinct parallel arcs. In such cases, an arc can be thought of as an ordered pair of vertices.



Example 1.7 We have a digraph with vertex set $V = \{v_1, v_2, v_3, v_4\}$ and arc set $A = \{a_1, a_2, \ldots, a_9\}$. The incidence function returns:

$$\begin{array}{llll}
\psi(a_1) & = (v_1, v_1) & \qquad \psi(a_6) & = (v_2, v_4) \\
\psi(a_2) & = (v_1, v_2) & \qquad \psi(a_7) & = (v_3, v_3) \\
\psi(a_3) & = (v_3, v_2) & \qquad \psi(a_8) & = (v_3, v_4) \\
\psi(a_4) & = (v_2, v_3) & \qquad \psi(a_9) & = (v_4, v_3) \\
\psi(a_5) & = (v_2, v_3) & &
\end{array}$$

Note that we can write shortly $a_1 = (v_1, v_1)$, $a_2 = (v_1, v_2)$, etc.

1.3 Graphs as Models

Graphs and digraphs have numerous applications in science, engineering, and the social sciences. Below, we give just a few examples of graph models that can be used to represent social, biological, communication, and information networks.

- **Niche overlap graphs in ecology**: $G = (V, E)$ where

  $V = \{\text{species in an ecosystem}\}$

  $uv \in E \quad \Leftrightarrow \quad$ species u and v compete for resources

- **Predator-prey graphs**: $D = (V, A)$ where

  $V = \{\text{species in an ecosystem}\}$

  $(u, v) \in A \quad \Leftrightarrow \quad$ species u preys on species v

- **Protein-protein interaction networks**: $G = (V, E)$ where

  $V = \{\text{proteins in a cell}\}$

  $uv \in E \quad \Leftrightarrow \quad$ proteins u and v interact

- **Friendship graphs**: $G = (V, E)$ where

  $V = \{\text{people in a group}\}$

  $uv \in E \quad \Leftrightarrow \quad$ persons u and v are friends

- **Mathematicians' collaboration networks**: $G = (V, E)$ where

  $V = \{\text{mathematicians}\}$

  each $uv \in E$ represents an article co-authored by mathematicians u and v

- **Tournament graphs**: $D = (V, A)$ where

  $V = \{\text{participants in a round-robin tournament}\}$

  $(u, v) \in A \quad \Leftrightarrow \quad$ participant u won a match against participant v

- **The web graph** $D = (V, A)$ where

  $V = \{\text{web pages}\}$

  $(u, v) \in A \quad \Leftrightarrow \quad$ web page u has a link to web page v

- **Transportation networks** (e.g. airline connection networks): $D = (V, A)$ where

  $V = \{\text{airports}\}$

  each $(u, v) \in A$ represents a direct flight from airport u to airport v (e.g., per day)

- **Call graphs**: $D = (V, A)$ where

  $V = \{\text{phone numbers}\}$

  each $(u, v) \in A$ represents a call from phone number u to phone number v (e.g., in a certain month)

In the rest of these notes, we shall focus on undirected graphs. However, many results for undirected graphs can be generalized to directed graphs.

1.4 Exercises

1. The **intersection graph** of a collection of sets $\{A_1, \ldots, A_n\}$ is the graph with vertices $A_1, \ldots, A_n$, where vertices A_i and A_j are adjacent if and only if their intersection is non-empty.

   Construct the intersection graph for each of the following collections of sets.

   (a) $A_1 = \{0, 2, 4, 6, 8\}$, $A_2 = \{0, 1, 2, 3, 4\}$, $A_3 = \{1, 3, 5, 7, 9\}$, $A_4 = \{5, 6, 7, 8, 9\}$, $A_5 = \{0, 1, 8, 9\}$

   (b) $A_1 = \{\ldots, -4, -3, -2, -1, 0\}$, $A_2 = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, $A_3 = \{\ldots, -6, -4, -2, 0, 2, 4, 6, \ldots\}$, $A_4 = \{\ldots, -5, -3, -1, 1, 3, 5, \ldots\}$, $A_5 = \{\ldots, -9, -6, -3, 0, 3, 6, 9, \ldots\}$

   (c) $A_1 = (-\infty, 0)$, $A_2 = (-1, 0)$, $A_3 = (0, 1)$, $A_4 = (-1, \infty)$, $A_5 = \mathbb{R}$

2. Let $\mathcal{R}$ be a binary relation on a set V. Define a digraph $D = (V, A)$ by

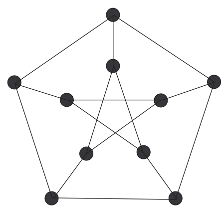
$$(x, y) \in A \quad \Leftrightarrow \quad x\mathcal{R}y.$$

   What can you say about the digraph D if the relation $\mathcal{R}$ is (a) symmetric; (b) reflexive; (c) irreflexive, (d) antisymmetric? Can D have multiple arcs?

3. The **Petersen graph** P can be defined as follows. Its vertices are the 2-subsets of $\{0, 1, 2, 3, 4\}$, and two subsets S_1 and S_2 are endpoints of an edge in P if and only if $S_1 \cap S_2 = \emptyset$.

   (a) How many vertices does the Petersen graph have?

   (b) Show that the figure below represents the Petersen graph by correctly labeling its vertices as 2-subsets of $\{0, 1, 2, 3, 4\}$.



4. Construct a **precedence digraph** for the following commands. The vertices of the digraph will be the commands (value assignments) $C_1, \ldots, C_7$, and there will be an arc with initial vertex C_i and terminal vertex C_j if command C_i must be executed before command C_j.

   $C_1 : x = 0$
   $C_2 : x = x + 1$
   $C_3 : y = 2$
   $C_4 : z = y$

$C_5 : x = x + 2$
$C_6 : y = x + z$
$C_7 : z = 4$

5. Construct a graph $G = (V, E)$ as follows. The vertex set V is the set of six of your chosen classmates, including yourself. For each two classmates a and b, there will be an edge in G for each class that a and b are taking together this term. How many loops does this graph have?

6. In a round-robin tournament, the Cave Lions beat the Dodos, Newfoundland Wolves, and Flying Foxes; the Dodos beat the Newfoundland Wolves and Flying Foxes, and the Newfoundland Wolves beat the Flying Foxes. Model this round-robin tournament with a digraph.

7. A student in a Joint Honours BSc in Computer Science and Mathematics has to take the following MAT courses: 1320 (prerequisite for 1325), 1325 (prerequisite for 2122 and 2125), 1341 (prerequisite for 2141 and 2143), 1348, 2122, 2125, 2141, and 2143. Model this requirement with an appropriate graph.

2 Graph Terminology and Special Graphs

2.1 Terminology

Definition 2.1 Let $G = (V, E)$ be a graph.

Vertices $u, v \in V$ are called **adjacent** or **neighbours** in G if uv is an edge of G. This is denoted by $u \sim_G v$, where the subscript G may be omitted if the graph G is understood.

An edge uv is said to be **incident** with each of its endpoints u and v.

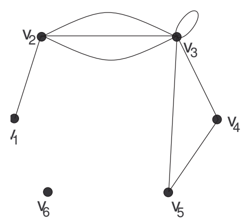
The **degree** of a vertex $u \in V$, denoted by $\deg_G(u)$, is the number of edges of G incident with vertex u, *each loop counting twice*.

A vertex of degree 0 is called **isolated**, and a vertex of degree 1 is called **pendant** (or a **leaf** in the context of trees).

If $V = \{v_1, v_2, \ldots, v_n\}$, then the sequence

$$(\deg_G(v_1), \deg_G(v_2), \ldots, \deg_G(v_n))$$

is called a **degree sequence** of G.



Example 2.2 The figure above represents a graph $G = (V, E)$ with $V = \{v_1, \ldots, v_6\}$ and degree sequence $(0, 1, 2, 2, 4, 7)$. (Note that the entries in a degree sequence may be listed in any order, though often they are listed in a non-decreasing order). Vertex v_6 is isolated in G, and vertex v_1 is pendant. Vertex v_3 has vertices v_2, v_4, v_5, as well as itself as a neighbour.

Theorem 2.3 The Handshaking Theorem
In any graph $G = (V, E)$, we have

$$\sum_{u \in V} \deg_G(u) = 2|E|.$$

Proof. In the sum $\sum_{u \in V} \deg_G(u)$:

- each loop incident with vertex u counts 2 towards $\deg_G(u)$, and
- each link uv counts 1 towards $\deg_G(u)$ and 1 towards $\deg_G(v)$.

Hence each edge is counted twice in $\sum_{u \in V} \deg_G(u)$, and the result follows. □

Exercise 2.4 How many edges in a graph with 5 vertices and degree sequence (2,2,3,3,4)?
Solution: Let $G = (V, E)$ be a graph with degree sequence (2,2,3,3,4). By the Handshaking Theorem,

$$2 + 2 + 3 + 3 + 4 = 2|E|,$$

whence $|E| = 7$.

Exercise 2.5 Suppose we have a graph $G = (V, E)$, in which every vertex has degree 3, 5, or 7. If G has 14 vertices and 29 edges, and has twice as many vertices of degree 3 as vertices of degree 5, how many vertices of each degree does it have?
Solution: Let x, y, and z be the numbers of vertices of degree 3, 5, and 7, respectively. By the Handshaking Theorem

$$3x + 5y + 7z = 2|E| = 58.$$

In addition, we know that $x + y + z = 14$ and $x = 2y$. We thus have a system of linear equations

$$\begin{aligned} 3x + 5y + 7z &= 58 \\ x + y + z &= 14 \\ x - 2y &= 0 \end{aligned}$$

with the unique solution $x = 8$, $y = 4$, and $z = 2$. Hence G has eight vertices of degree 3, four vertices of degree 5, and two vertices of degree 7.

Corollary 2.6 Every graph has an even number of vertices of odd degree.

Proof. Let $G = (V, E)$ be a graph, V_1 the subset of V containing all vertices of odd degree, and $V_2 = V - V_1$. Then, by the Handshaking Theorem,

$$2|E| = \sum_{u \in V} \deg_G(u) = \sum_{u \in V_1} \deg_G(u) + \sum_{u \in V_2} \deg_G(u),$$

so

$$\sum_{u \in V_1} \deg_G(u) = 2|E| - \sum_{u \in V_2} \deg_G(u).$$

On the right hand side, we have a sum of even numbers, so $\sum_{u \in V_1} \deg_G(u)$ must be even. But each of the $|V_1|$ terms in this sum is odd, whence $|V_1|$ must be even. $\square$

It can be shown that any sequence of non-negative integers containing an even number of odd entries is a degree sequence of a graph, but not necessarily of a simple graph. Some examples can be seen below.

Exercise 2.7 Does there exist a graph (simple graph) with the following degree sequence? If so, draw a picture of such a graph. If not, explain why.

1. $(1, 2, 2, 4, 5, 5)$

   Solution: This sequence contains an odd number of odd entries, hence by Corollary 2.6 it cannot be a degree sequence of any graph.

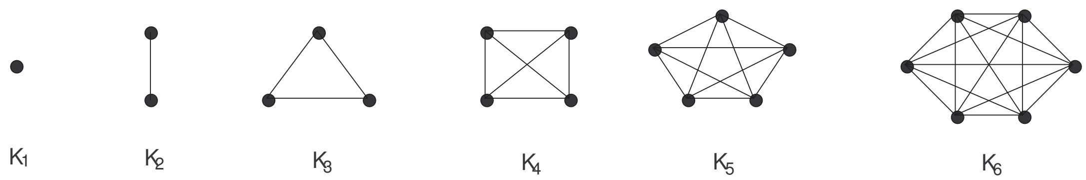
2. $(3, 3, 3, 3, 3, 5)$

   Solution: There exists a simple graph with this degree sequence (see the figure below, left).

3. $(0, 2, 2, 3, 4, 5)$

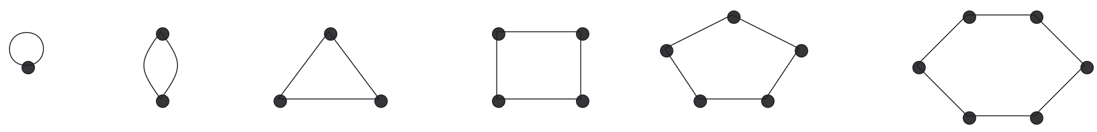
   Solution: The figure above (right) shows a graph with this degree sequence. However, we claim that there is no simple graph with degree sequence $(0, 2, 2, 3, 4, 5)$. To the contrary, suppose that $G = (V, E)$ is a simple graph with this degree sequence. Then G has 6 vertices, including vertices u and v such that $\deg(u) = 0$ and $\deg(v) = 5$. The latter means that v is adjacent to all other vertices in G (including u), while the former means that u is adjacent to none. This is a contradiction. Hence no simple graph with this degree sequence exists.

2.2 Some Special Graphs



- A **complete graph** K_n (for $n \geq 1$) is a simple graph with n vertices in which every pair of distinct vertices are adjacent. More formally,

$$\begin{aligned} V(K_n) &= \{u_1, u_2, \ldots, u_n\} \\ E(K_n) &= \{xy : x, y \in V, x \neq y\} \end{aligned}$$

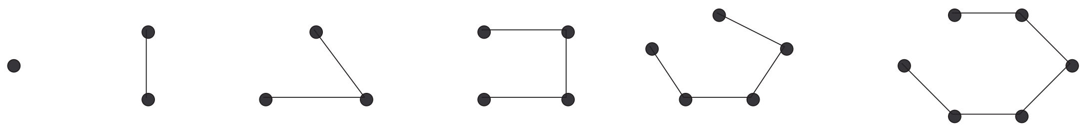


C_1 C_2 C_3 C_4 C_5 C_6

- A **cycle** C_n (of **length** $n \geq 1$) is a graph with n vertices that are linked in a circular way, creating n edges. That is,

$$\begin{aligned} V(C_n) &= \{u_1, u_2, \ldots, u_n\} \\ E(C_n) &= \{u_1u_2, u_2u_3, u_3u_4, \ldots, u_{n-1}u_n, u_nu_1\} \end{aligned}$$

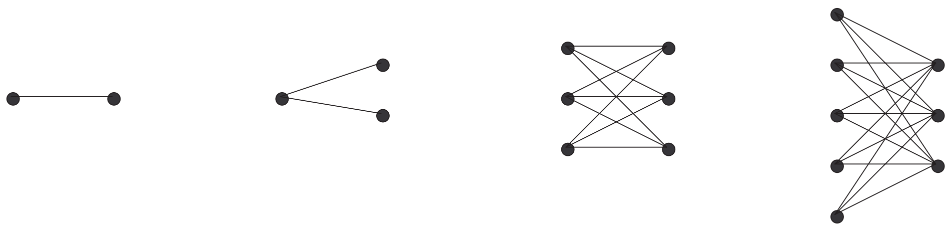
  Note that any cycle C_n for $n \geq 3$ is a simple graph, while for $n = 2$, the edge set $E(C_n)$ consists of a pair of parallel edges.



P_0 P_1 P_2 P_3 P_4 P_5

- A **path** P_n (of **length** $n \geq 0$) is a simple graph with $n+1$ vertices that are linked in a linear way, creating n edges. More precisely,

$$\begin{aligned} V(P_n) &= \{u_0, u_1, u_2, \ldots, u_n\} \\ E(P_n) &= \{u_0u_1, u_1u_2, u_2u_3, \ldots, u_{n-1}u_n\} \end{aligned}$$



$K_{1,1}$ $K_{1,2}$ $K_{3,3}$ $K_{5,3}$

- A **complete bipartite graph** $K_{m,n}$ (for $m, n \geq 1$) is a simple graph with $m+n$ vertices. The vertex set partitions into sets X and Y of cardinalities m and n, and each pair of vertices from distinct parts are adjacent. That is,

$$\begin{aligned} V(K_{m,n}) &= \{x_1, x_2, \ldots, x_m\} \cup \{y_1, y_2, \ldots, y_n\} \\ E(K_{m,n}) &= \{x_iy_j : x_i \in X, y_j \in Y\} \end{aligned}$$

2.3 Subgraphs

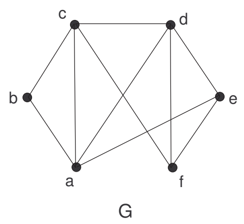
Example 2.8 Consider the two graphs below. Observe that

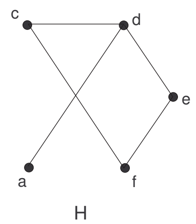
$$V(H) = \{a, c, d, e, f\} \subseteq \{a, b, c, d, e, f\} = V(G)$$

and

$$E(H) = \{ad, cd, cf, de, ef\} \subseteq \{ab, ac, ad, ae, bc, cd, cf, de, df, ef\} = E(G).$$

Thus H is a graph within the graph G. We say that H is a *subgraph* of G. A general definition follows below.





Definition 2.9 Let G and H be simple graphs. We say that H is a **subgraph** of G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$.

This definition can be extended to non-simple graphs, however, an additional requirement on the incidence functions is needed to make it precise. Namely, the incidence function of H must be equal to the *restriction* of the incidence function of G to the set $E(H)$; in symbols, $\psi_H = \psi_G|_{E(H)}$.

Exercise 2.10 Let G be a graph with vertex set $V = \{a, b, c, d\}$ and edge set $E = \{ab, ac, ad, bc\}$. Find all subgraphs of the graph G with exactly 3 vertices and 2 edges.

Solution: There are $\binom{4}{3} = 4$ subsets of V of cardinality 3, namely

$$V_1 = \{a, b, c\}, \quad V_2 = \{a, b, d\}, \quad V_3 = \{a, c, d\}, \quad \text{and} \quad V_4 = \{b, c, d\}.$$

The subgraphs with vertex set V_1 and exactly two edges have, respectively, edge sets

$$E_1 = \{ab, ac\}, \quad E_1' = \{ab, bc\}, \quad \text{and } E_1'' = \{ac, bc\};$$

the one subgraph with vertex set V_2 and exactly two edges has edge set

$$E_2 = \{ab, ad\};$$

and the one subgraph with vertex set V_3 and exactly two edges has edge set

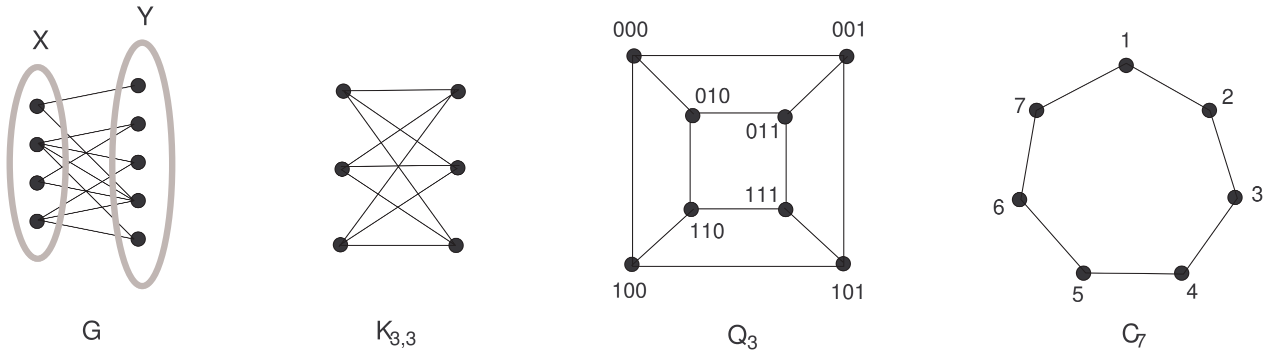
$$E_2 = \{ac, ad\}.$$

There is no subgraph with vertex set V_4 and exactly two edges.

2.4 Bipartite Graphs

We now generalize complete bipartite graphs as follows.

Definition 2.11 A graph $G = (V, E)$ is called **bipartite** if V can be partitioned into two subsets X and Y such that every edge of G has one endpoint in X and the other endpoint in Y. If this is the case, then $\{X, Y\}$ is called a **bipartition** of G, and subsets X and Y are the two **parts**. (See the figure below, extreme left.)



Exercise 2.12 Are the graphs $K_{3,3}$, Q_3, and C_7 in the figure above bipartite? If the graph is bipartite, find a bipartition; otherwise, explain why it is not bipartite.

Solution: The graph $K_{3,3}$ is clearly bipartite; the vertices on the left form one part, and the vertices on the right the other part in the bipartition. In fact, any complete bipartite graph $K_{m,n}$ is bipartite.

The graph Q_3, the 3-dimensional cube, can be seen to be bipartite as well. If we put vertex 000 into part X, all its neighbours into part Y, all their neighbours into part X, etc., we end up with $X = \{000, 011, 101, 110\}$ and $Y = \{001, 010, 100, 111\}$, and since every edge joins vertices that differ in exactly one bit, every edge has one endpoint in X and the other in Y. We conclude that $\{X, Y\}$ is a bipartition for Q_3, and Q_3 is bipartite.

The same strategy applied to the cycle C_7 fails: without loss of generality, put vertex 1 into part X, then vertex 2 has to go into Y, vertex 3 into X again, and so on. Vertex 7 ends up in part X, and we have the edge 17 with both ends in part X. Hence C_7 is not bipartite. This observation extends to all cycles of odd length: odd cycles are not bipartite. In fact, an odd cycle is the most basic structure that prevents a graph from being bipartite.

Theorem 2.13 The following statements about a graph G are equivalent:

1. G is bipartite;
2. G admits a proper 2-vertex-colouring; that is, the vertices of G can be coloured with 2 colours (say, red and blue) so that the endpoints of each edge receive distinct colours;
3. G has no subgraph that is a cycle of odd length.

Proof. (1) $\Rightarrow$ (2): Assume G is bipartite with bipartition $\{X, Y\}$. Colour the vertices in X red and the vertices in Y blue. Since each edge has one endpoint in X and the other in Y, the endpoints of each edge receive distinct colours. We conclude that G admits a proper 2-vertex-colouring.

(2) $\Rightarrow$ (1): Assume that G admits a proper 2-vertex-colouring. Let X be the set of all vertices in $V(G)$ coloured red, and Y be the set of all vertices in $V(G)$ coloured blue. Since the endpoints of each edge receive distinct colours, each edge must have one endpoint in X and the other in Y. We conclude that G is bipartite with bipartition $\{X, Y\}$.

(2) $\Rightarrow$ (3): Assume that G admits a proper 2-vertex-colouring, and suppose that it has a subgraph C that is an odd cycle. Then the vertices in C must alternate in colour, which means that half of the vertices are blue and half of them are red. A contradiction, since C has an odd number of vertices. Hence G has no subgraph that is an odd cycle.

(3) $\Rightarrow$ (2): This part of the proof is more challenging, and will be omitted. $\square$

2.5 Exercises

1. How many edges does the following graph have:
   (a) a graph with degree sequence $(0, 0, 1, 1, 2, 3, 4, 5, 5, 7)$?
   (b) a graph with 12 vertices, of which half have degree 3 and the others degree 4?
2. If G is a graph with 11 vertices and 20 edges such that every vertex has degree 3 or 4, how many vertices of degree 3 does G have?
3. If G is a simple graph with 6 vertices and 10 edges in which every vertex has odd degree, and the number of vertices of degree 3 is one more that the number of vertices of degree 5, how many vertices of each degree does G have?
4. For each of the special graphs K_n, C_n, P_n, and $K_{m,n}$, determine the degree sequence.
5. A graph is called **regular of degree** k (shortly k**-regular**) if every vertex has degree k. For what values of the parameters are the graphs K_n, C_n, P_n, and $K_{m,n}$ regular?
6. (a) Let G be a simple graph. Show that G has two vertices of the same degree.

   (b) Show that in any group of people, there are two people with the same number of friends within the group (not counting themselves). *Hint: use part (a).*
7. For what values of the parameters are the special graphs K_n, C_n, P_n, and $K_{m,n}$ bipartite?
8. The **wheel** graph, denoted W_n, is the simple graph obtained from the cycle C_n by adjoining a new vertex and joining it with an edge to each of the vertices of the cycle.
   (a) Draw the first five wheels W_i for $i = 3, 4, \ldots, 7$.
   (b) Determine the degree sequence of the graph W_n.
   (c) For what values of the parameter n is the wheel W_n regular?
   (d) For what values of the parameter n is the wheel W_n bipartite?
9. Does there exist (a) a graph, (b) a simple graph, with the following degree sequence? If so, draw a picture of such a graph. If not, explain why not.

(a) $(2, 5, 5, 5, 5, 5, 5)$

(b) $(3, 3, 3, 3, 3, 5, 5)$

(c) $(2, 2, 2, 3, 4, 4, 7)$

(d) $(0, 2, 2, 3, 4, 5, 6)$

10. (a) Show that any sequence $(d_1, d_2, \ldots, d_n)$ of non-negative even integers is a degree sequence of a graph (not necessarily simple). *Hint: use loops.*

    (b) Show that any sequence $(d_1, d_2, \ldots, d_n)$ of non-negative integers such that $\sum_{i=1}^{n} d_i$ is even is a degree sequence of a graph (not necessarily simple). *Hint: use (a).*

11. List all subgraphs of the following graphs: (a) K_2, (b) K_3, (c) C_4, (d) P_2, (e) W_3.

12. The k**-cube** Q_k (for $k = 1, 2, \ldots$) is the graph whose vertices are the binary strings of length k, and two vertices are adjacent if and only if they differ in exactly one bit (e.g. for $k = 5$, vertices 01001 and 11001 are adjacent, while 01001 and 00011 are not).

    (a) Draw the cubes Q_k for $k = 1, 2, 3$.

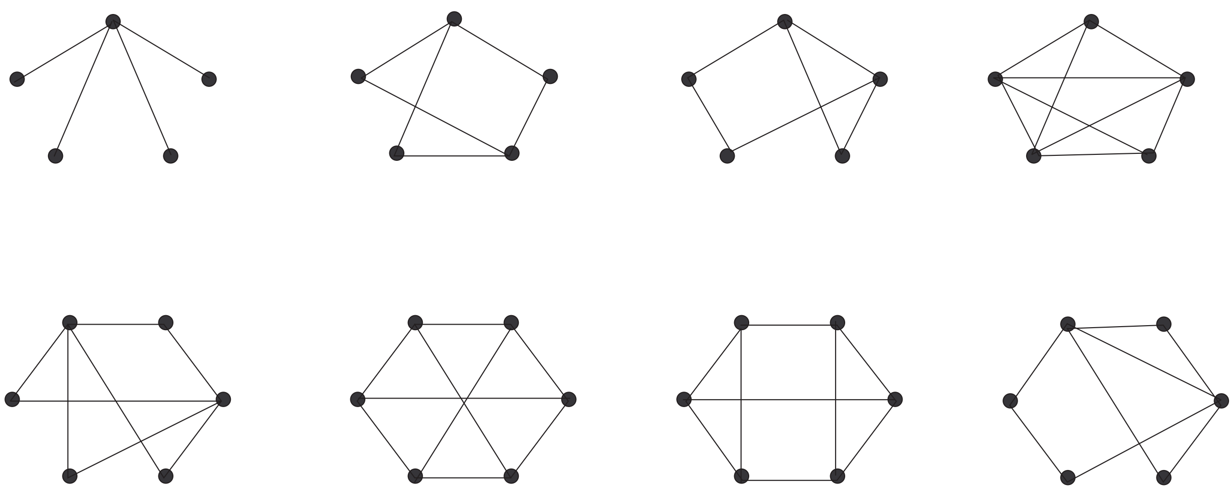
    (b) Show that Q_k has 2^k vertices.

    (c) Show that Q_k is k-regular.

    (d) Show that Q_k has $k2^{k-1}$ edges.

    (e) Show that Q_k is bipartite.

13. For each of the graphs in the figure, determine whether or not it is bipartite. If the graph is bipartite, give a proper 2-vertex colouring. If not, find a subgraph that is an odd cycle.



3 Matrices Associated with a Graph, and Graph Isomorphism

3.1 Matrices Associated with a Graph

Definition 3.1 Let G be a graph with $V(G) = \{u_1, u_2, \ldots, u_n\}$, $E(G) = \{e_1, e_2, \ldots, e_m\}$, and incidence function ψ_G. We define:

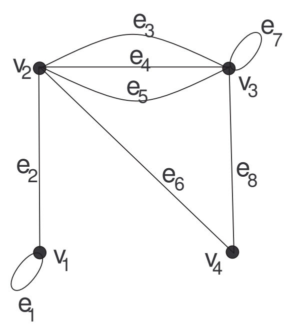
- the **incidence matrix** of G: an $n \times m$ matrix $M = [m_{ij}]$ such that

$$m_{ij} = \begin{cases} 2 & \text{if } \psi_G(e_j) = \{u_i\} \\ 1 & \text{if } \psi_G(e_j) = \{u_i, u_k\} \text{ for some } k \neq i \\ 0 & \text{otherwise} \end{cases}$$

- the **adjacency matrix** of G: an $n \times n$ matrix $A = [a_{ij}]$ such that

$$a_{ij} = |\{e_k : \psi_G(e_k) = \{u_i, u_j\}\}| = \text{ number of edges with endpoints } u_i \text{ and } u_j.$$

Note: Strictly speaking, it is incorrect to refer to *the* incidence matrix of a graph because each ordering of the vertices and edges may give a different matrix. However, once the ordering is specified, the incidence matrix is indeed unique. Similarly, the adjacency matrix is unique once the ordering of the vertices has been specified.



Example 3.2 Find the incidence matrix and the adjacency matrix of the graph in the figure above.

Solution: We have

$$M = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 3 & 1 \\ 0 & 3 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

Exercise 3.3 Let G be a graph with $V(G) = \{u_1, u_2, \ldots, u_n\}$ and $E(G) = \{e_1, e_2, \ldots, e_m\}$, M its incidence matrix, and A its adjacency matrix.

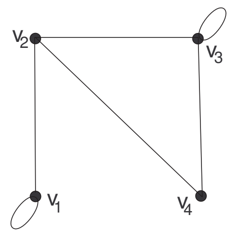
1. What are the row and column sums in M?
2. If G is simple, what are the row and column sums in A? What other properties does A have?

Solution:

1. In the incidence matrix M, each column sum is 2. Row i corresponds to the vertex u_i; in this row, each entry 2 corresponds to a loop incident with u_i, and each entry 1 corresponds to a link incident with u_i. Hence the sum of the entries in row i is $\deg(u_i)$.

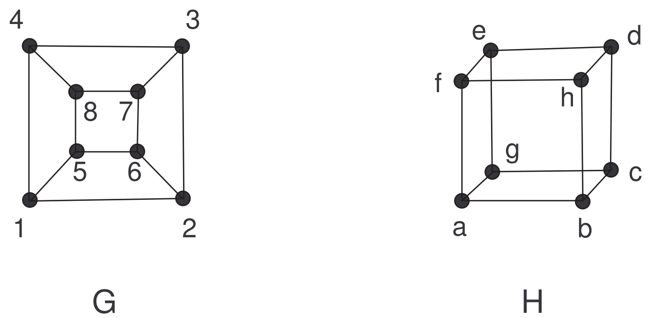
2. In the adjacency matrix A, row i and column i both correspond to vertex u_i. In fact, A is a symmetric matrix, and if G is simple, then each entry in A is either 0 or 1, and all diagonal entries are 0. The sum of row i is then equal to the number of edges incident with u_i, which is (for simple graphs) $\deg(u_i)$.

Another way to represent a graph without multiple edges is using **adjacency lists**: for each vertex v of the graph, the adjacency list of v is the list of the neighbours of vertex v in G.



Example 3.4 In the above example, we have the following adjacency lists.
Vertex v_1: v_1, v_2
Vertex v_2: v_1, v_3, v_4
Vertex v_3: v_2, v_3, v_4
Vertex v_4: v_2, v_3

3.2 Graph Isomorphism



Example 3.5 Consider the two graphs in the figure. They both represent the skeleton of a cube, so they are "essentially the same", just drawn from a different perspective and having

different vertex labels. In fact, we can relabel the vertices so that the relabelling function preserves adjacency. For example, if $\varphi : V(G) \to V(H)$ is defined as

$$\begin{aligned} \varphi(1) &= a & \varphi(5) &= f \\ \varphi(2) &= b & \varphi(6) &= h \\ \varphi(3) &= c & \varphi(7) &= d \\ \varphi(4) &= g & \varphi(8) &= e \end{aligned}$$

then φ maps any pair of adjacent vertices in G to a pair of adjacent vertices in H, and conversely. Observe also that φ is a bijection. Such a relabelling function is called an isomorphism.

Definition 3.6 Let G and H be simple graphs. An **isomorphism** from G to H is a bijection $\varphi : V(G) \to V(H)$ such that

$$u \sim_G v \quad \Leftrightarrow \quad \varphi(u) \sim_H \varphi(v)$$

for all $u, v \in V(G)$. Graphs G and H are called **isomorphic** (denoted $G \cong H$) if there exists an isomorphism from G to H.

A **graph invariant** is a graph property or parameter that is preserved under isomorphisms; that is, isomorphic graphs must agree on this property or parameter. Many graph properties are invariants; for example:

- number of vertices,
- number of edges,
- degree sequence,
- being bipartite,
- containing specific subgraphs,

and many more. In fact, all properties that are purely structural are invariants. So what properties are not invariants? Anything that is separate from and added to the structure, for example

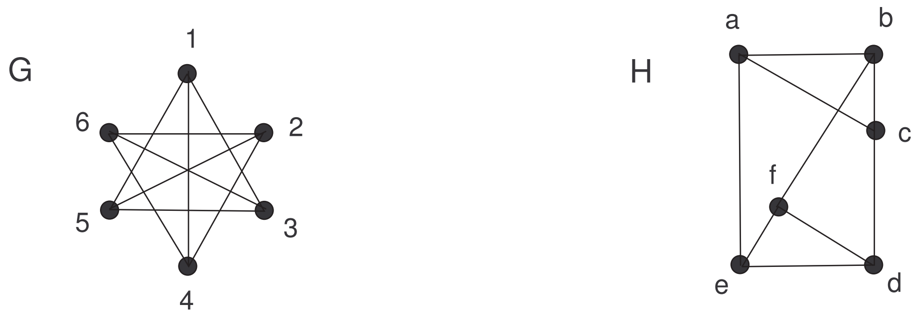
- vertex and edge labels,
- vertex and edge colours,
- drawing.

Important observation:

- To prove that graphs G and H are isomorphic, we must *find an isomorphism* from G to H.
- To prove that graphs G and H are *not* isomorphic, it suffices to *find an invariant* in which G and H differ.

Example 3.7 Are the following pairs of graphs isomorphic?

Solution: No, these two graphs are not isomorphic. H has a subgraph isomorphic to C_3, while G does not (it is, in fact, bipartite and isomorphic to $K_{3,3}$).



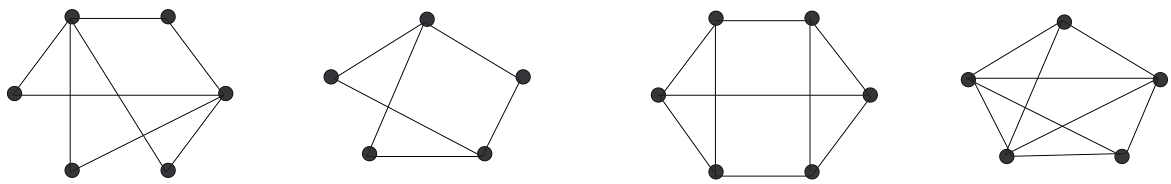
Solution: Yes, these two graphs are isomorphic. An isomorphism $\varphi : V(G) \to V(H)$ is given by

$$\begin{aligned} \varphi(1) &= a & \varphi(4) &= e \\ \varphi(2) &= d & \varphi(5) &= c \\ \varphi(3) &= b & \varphi(6) &= f \end{aligned}$$

Solution: No, these two graphs are not isomorphic. H has a subgraph isomorphic to C_8, while G does not. The latter can be shown by contradiction: a C_8 would have to contain exactly two edges incident with each vertex; an attempt to construct such a cycle in G quickly leads to a contradiction.

Solution: No, these two graphs are not isomorphic. They have the same degree sequence, however, graph H contains no pair of adjacent vertices of degree 3, while G does.

3.3 Exercises



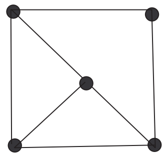
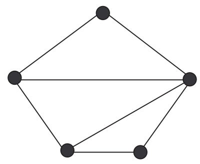
1. For each of the graphs drawn above, find an incidence matrix and an adjacency matrix.
2. (a) For each matrix B_i below, draw a graph whose (i) adjacency matrix, (ii) incidence matrix is B_i.

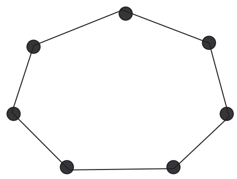
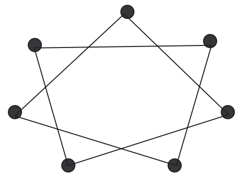
$$B_1 = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}, \qquad B_2 = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

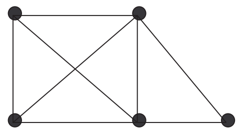
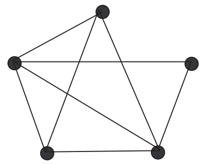
   (b) Represent each of the graphs from (ii) using adjacency lists.
3. Draw a graph with adjacency matrix

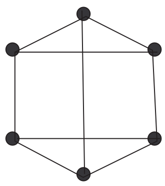
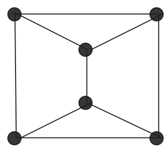
$$A_1 = \begin{bmatrix} 1 & 3 & 2 \\ 3 & 0 & 4 \\ 2 & 4 & 0 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 2 & 0 & 3 & 0 \\ 0 & 3 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

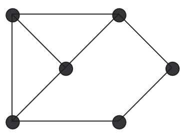
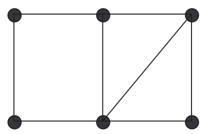
4. Find an adjacency matrix and an incidence matrix for each of the following graphs: K_4, $K_{1,4}$, $K_{2,3}$, C_4, P_5.
5. Describe the adjacency matrix and the incidence matrix for each of the following graphs: K_n, $K_{m,n}$, C_n.
6. Find all (pairwise non-isomorphic) graphs with 3 vertices whose adjacency and incidence matrix are the same.
7. For each pair of simple graphs G and H below, determine whether the graphs are isomorphic or not. Justify your answer: if you claim that the graphs are isomorphic, give an isomorphism; if you claim that they are not, give an invariant in which they differ.

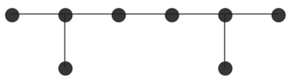
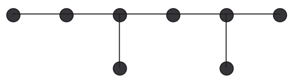
(a) G  H 

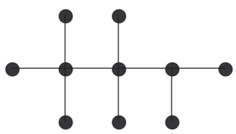
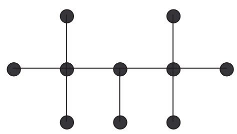
(b) G  H 

(c) G  H 

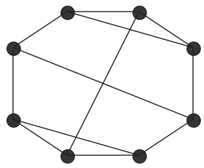
(d) G  H 

(e) G  H 

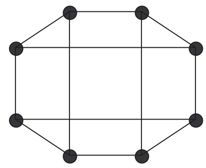
(f) G  H 

(g) G  H 

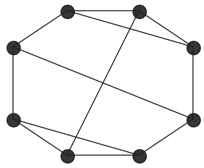
(i) G



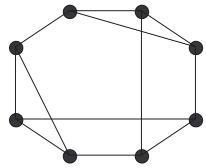
H



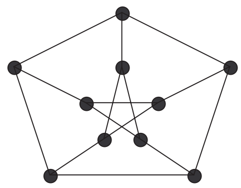
(j) G



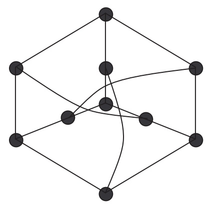
H



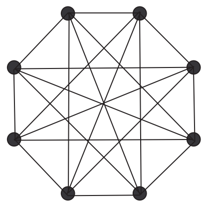
(k) G



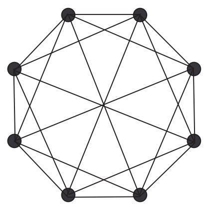
H



(l) G



H



8. How many pairwise non-isomorphic simple graphs are there with

   (a) n vertices, for $n = 2, 3, 4$?

   (b) 5 vertices and 4 edges?

   (c) 6 vertices and 4 edges?

9. Describe how you would determine whether or not graphs G and H are isomorphic by comparing their adjacency matrices.

10. Show that isomorphism of simple graphs is an equivalence relation on the set of all simple graphs.

4 Walks and Connection

4.1 Walks, Trails, Paths, Cycles

Definition 4.1 Let $G = (V, E)$ be a graph with the incidence function ψ_G. Let $x, y \in V$ and $k \in \mathbb{N}$. An (x, y)**-walk of length** k in G is an alternating sequence of vertices and edges

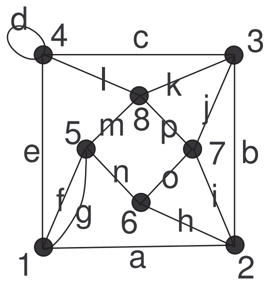
$$W = v_0 e_1 v_1 e_2 v_2 \dots v_{k-1} e_k v_k$$

such that

- $v_0, v_1, \dots, v_k \in V$,
- $e_1, e_2, \dots, e_k \in E$,
- $v_0 = x$ and $v_k = y$, and
- $\psi_G(e_i) = v_{i-1}v_i$ for all $i = 1, 2, \dots, k$.

A walk $W = v_0 e_1 v_1 e_2 v_2 \dots v_{k-1} e_k v_k$ is called

- **closed** if $v_0 = v_k$, and **open** otherwise;
- a **trail** if its edges are pairwise distinct;
- a **path** if its vertices are pairwise distinct; and
- a **cycle** if $v_0 = v_k$ while its **internal** vertices $v_1, \dots, v_k$ are pairwise distinct.



Example 4.2 In the graph above, find the following.

1. A (1,3)-walk of length 6 that is not a trail:

$$W = 1a2h6n5f1a2b3$$

2. A (1,3)-trail of length 3 that is not a path:

$$W = 1e4d4c3$$

3. A (1,3)-path of length 3:

$$W = 1f5m8k3$$

4. A closed walk of length 6 that contains vertex 2 and is not a trail:

$$W = 2b3j7i2h6o7i2$$

5. A closed trail of length 4 that contains vertex 4 and is not a cycle:

$$W = 4d4l8k3c4$$

6. A cycle of length 4 containing vertex 2:

$$W = 2b3j7o6h2$$

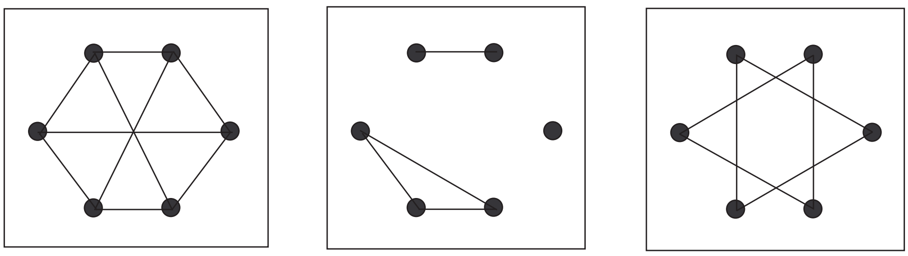
Note: In a simple graph, a walk $W = v_0e_1v_1e_2v_2 \dots v_{k-1}e_kv_k$ can be more simply denoted by its sequence of vertices, that is, as $W = v_0v_1v_2 \dots v_{k-1}v_k$

4.2 Connected Graphs

Definition 4.3 A graph $G = (V, E)$ is called **connected** if for any $x, y \in V$ there exists an (x, y)-path (or equivalently, (x, y)-walk) in G.

A graph that is not connected is called **disconnected**.

Note: It can be shown that in any graph G, for any vertices x and y, there exists an (x, y)-path in G of and only if there exists an (x, y)-walk.



Example 4.4 Of the graphs above only the first one is connected.

Definition 4.5 Maximal connected subgraphs of a graph G are called the **connected components** of G.

Notes:

- A connected subgraph H of a graph G is **maximal** (with respect to connectedness) if it is not contained in any other connected subgraph of G; that is, if adjoining any new vertex or edge of G to H would result in a disconnected subgraph of G.
- Every graph is a **vertex-disjoint union** of its connected components.
- A graph is connected if and only if it has exactly one connected component; namely, itself.

Example 4.6 The graphs above have 1, 3, and 2 connected components, in this order.

4.3 Exercises

1. What can you say about vertices x and y in the following graph models if you know that there exists an (x, y)-path? What do the connected components in these graph models represent?

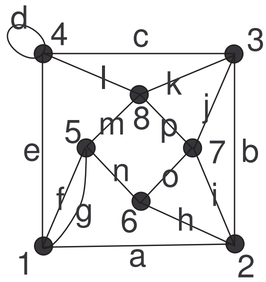
   (a) a niche overlap graph

   (b) a protein-protein interaction network

   (c) a friendship graph

   (d) a mathematicians' collaboration network

   (e) an airline connection network (as an undirected graph)



2. In the graph above, find the following walks, if they exist.

   (a) A (5,3)-walk of length 4 that is not a trail.

   (b) A (5,3)-trail of length 4 that is not a path.

   (c) A (5,3)-path of length 4.

   (d) A closed walk of length 4 that contains vertex 5 and is not a trail.

   (e) A closed trail of length 4 that contains vertex 5 and is not a cycle.

   (f) A cycle of length 4 containing vertex 5.

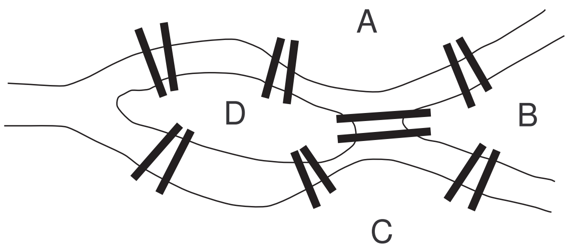
3. Draw all pairwise non-isomorphic simple graphs G with the following properties:

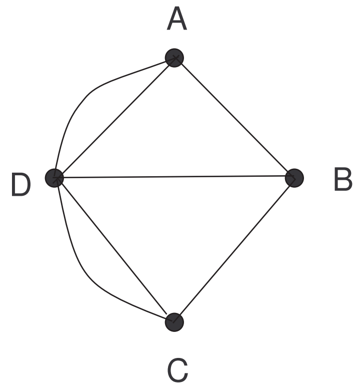
   (a) G is connected and has 4 vertices;

   (b) G has 5 vertices and exactly 2 connected components;

   (c) G is bipartite with 6 vertices and exactly 3 connected components.

5 Euler tours and trails



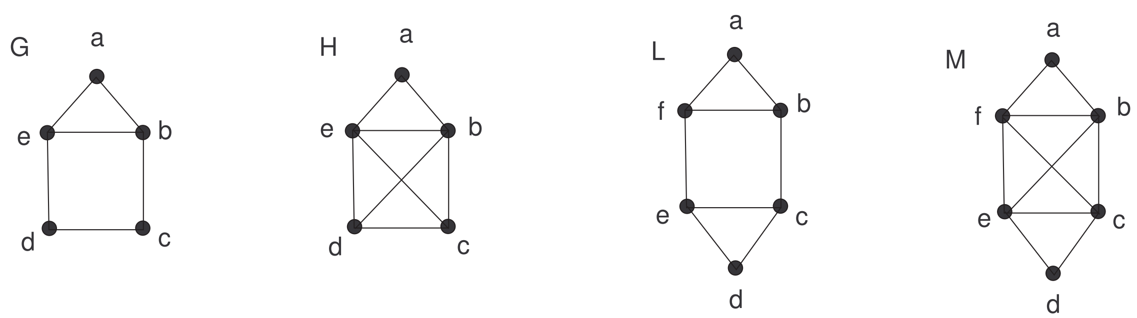


Example 5.1 The origins of graph theory — the Bridges of Königsberg

In 1736, while stationed in St. Petersburg, the Swiss mathematician Leonhard Euler took interest in the following puzzle. The townsfolk of Königsberg, Prussia (now Kaliningrad, Russia) take long Sunday walks. They wonder if it is possible to walk around the town, traverse each of the seven bridges of Königsberg exactly once and return to the starting point (see the map on the left).

This puzzle is equivalent to the question whether the graph on the right (representing the map of the town) admits a closed trail that traverses each edge of the graph (exactly once). In Euler's honour, such a closed trail is called an *Euler tour*.

Definition 5.2 An **Euler tour** in a graph G is a closed trail of G traversing each edge of G. An **Euler trail** in a graph G is an *open* trail of G traversing each edge of G.



Example 5.3 For each the graphs above, determine whether it has an Euler tour or an Euler trail.

Solution: Graph G admits an Euler trail, for example, $W = bcdeabc$, but no Euler tour.
Graph H admits an Euler trail, for example, $W = cdeabcebd$, but no Euler tour.
Graph L admits neither an Euler tour nor Euler trail.
Graph G admits an Euler tour, for example, $W = abcdefbecfa$, but no Euler trail.

The above examples suggest the following theorem.

Theorem 5.4 *Let G be a connected graph. Then:*

1. *G has an Euler tour if and only if it has no vertices of odd degree.*

2. *G has an Euler trail if and only if it has exactly two vertices of odd degree.*

PROOF. First note that any graph without isolated vertices that admits an Euler trail or Euler tour must be connected, since the trail or tour contains a walk between any pair of vertices of G. Hence let $G = (V, E)$ be a connected graph.

1. ($\Rightarrow$): Suppose G has an Euler tour T. Each time an *internal* vertex x is visited by the tour, 2 edges of the graph are traversed, contributing 2 to the degree of x. Hence $\deg_G(x)$ is even for all internal vertices of the tour. For the initial (= terminal) vertex of the tour (say u), the initial edge of the tour adds 1 to $\deg_G(u)$, each visit of u as an internal vertex of the tour adds 2, and the last edge of the tour adds 1 to $\deg_G(u)$. Hence $\deg_G(u)$ is also even. Therefore, if G admits an Euler tour, then every vertex of G has even degree.

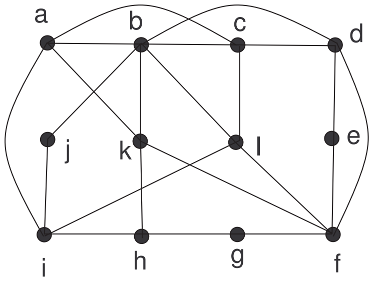
   ($\Leftarrow$): Suppose G has no vertices of odd degree. We construct an Euler tour in G as follows. Starting from any vertex, let T be a *longest* trail (that is, any trail that can not be extended). Since every vertex in G has even degree and T can not be extended, T must be a closed trail. If T traverses all edges of G, then T is the required Euler tour of G. Otherwise, remove the edges of T from G (as well as any isolated vertices) to obtain a graph G'. Note that the degrees of all vertices in G' are even. Since G is connected, T and G' must have a common vertex, say w. Now let T' be a longest trail in G' starting from vertex w. As above, T' must in fact be a closed trail. Join T and T' into a longer closed trail T'', using their common vertex w. (That is, we obtain T'' by traversing T up to vertex w, then traverse T' entirely from w, then continue along T back to its initial vertex.) Now replace T by T'', and continue the process until all edges of G have been used up. Since G has a finite number of edges, this process indeed terminates, producing an Euler tour of G.

2. ($\Rightarrow$): Suppose G has an Euler trail T. Counting as above, we can see that every internal vertex of the trail must have even degree in G, while the initial and the terminal vertex must have odd degree.

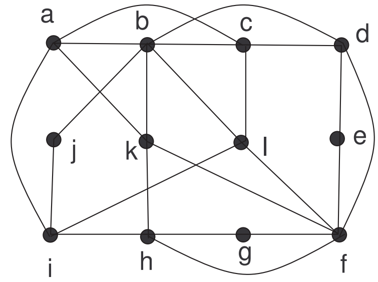
   ($\Leftarrow$): Suppose G has exactly two vertices of odd degree, say x and y. Obtain a graph G' by adjoining a new edge xy to G (there may be more than one edge with endpoints x and y in G'). Observe that G' has no vertices of odd degree since $\deg_{G'}(x) = \deg_G(x) + 1$, $\deg_{G'}(y) = \deg_G(y) + 1$, and $\deg_{G'}(u) = \deg_G(u)$ for all vertices $u \in V(G) - \{x, y\}$. The graph G' is also connected since G is. Hence G' admits an Euler tour T by (1). Removing the edge xy from T results in an Euler trail of G, necessarily with initial vertex x and terminal vertex y, or vice-versa.

$\square$

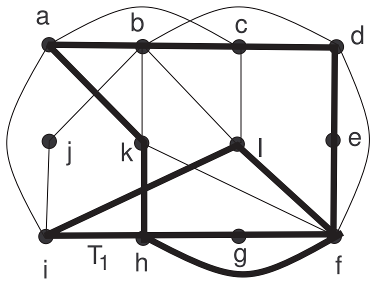
Example 5.5 Does the graph G below have an Euler tour? Does it have an Euler trail? If so, construct one with the help of the algorithm described in the proof of Theorem 5.4.



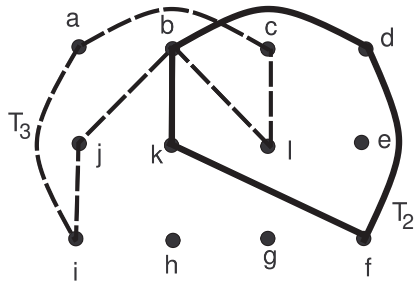
G



G+hf



G+hf



(G+hf)-E(T1)

Solution: G has exactly two vertices of odd degree, so it has no Euler tour, but it does have an Euler trail. We shall construct it using the algorithm of the proof of Theorem 5.4. First, we insert a new edge hf between the two odd-degree vertices of G, to obtain the graph $G + hf$. This graph has no vertices of odd degree (and it is connected), so it must admit an Euler tour. Starting at vertex h and edge hf, we first construct a longest trail

$$T_1 = hfedcbakhgflih.$$

It is necessarily a closed trail, ending at h. Then, after removing the edges of T_1, we construct a longest trail T_2 from vertex f, and after removing the edges of T_2, the remaining edges form a closed trail T_3, where

$$T_2 = fdbkf \qquad \text{and} \qquad T_3 = blcaijb.$$

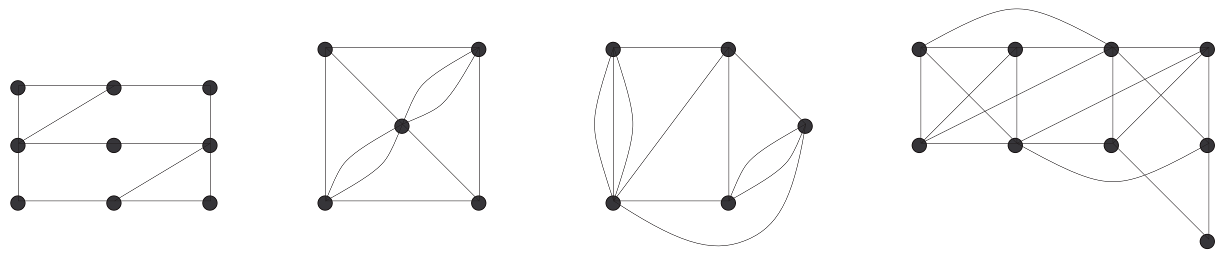
Next, we concatenate T_1, T_2, and T_3 into an Euler tour T of $G + hf$ using their intersecting vertices f and b, respectively, as points of connection:

$$T = hfdblcaijbkfedcbakhgflih.$$

Finally, removing the edge hf from T results in an Euler trail T' in G:

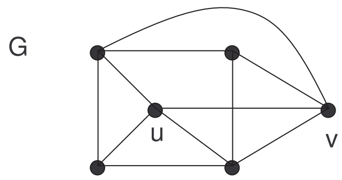
$$T' = fdblcaijbkfedcbakhgflih.$$

Note that T' necessarily starts and ends at a vertex of odd degree.



5.1 Exercises

1. For each of the (connected) graphs above, determine whether or not it has an Euler tour or an Euler trail. If so, construct one. (Use the algorithm described in the proof of Theorem 5.4 if needed.) In either case, justify your answer.



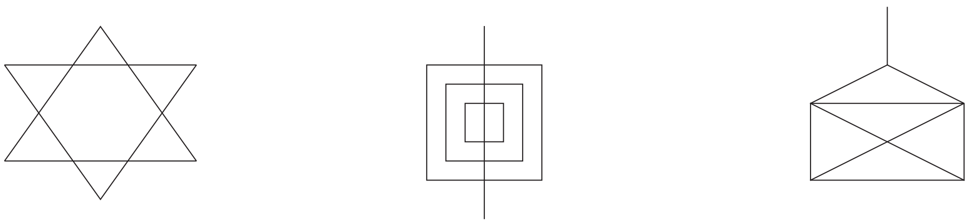
2. For any graph G and edge $uv \in E(G)$, the symbol $G - uv$ denotes the graph obtained from G by deleting one copy of the edge uv. Similarly, for any graph G and vertices $u, v \in V(G)$, the symbol $G + uv$ denotes the graph obtained from G by adjoining a new copy of the edge uv.

   For the graph G above, determine whether G, $G - uv$, and $G + uv$ admit an Euler tour or an Euler trail. Justify your answer.

3. For what values of the parameters do the following graphs admit (a) an Euler tour, (b) an Euler trail?

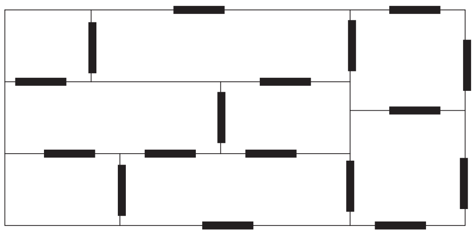
   (i) K_n, (ii) C_n, (iii) $K_{m,n}$, (iv) W_n, (v) Q_n

   In each of the questions below, model the problem with an appropriate graph. What relevant properties does this graph have? How does the original problem translate into a problem about graphs? Justify your answer by referring to the appropriate theorem.

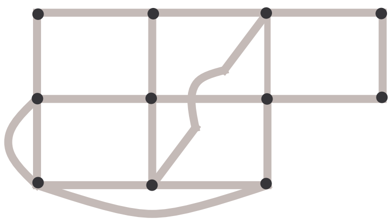


4. Can the three figures above be drawn with a pencil in a continuing motion without lifting the pencil or retracing a part of the picture?

5. Can the townsfolk of Königsberg walk through the town crossing each bridge exactly once and ending in a different part of town? What if two additional bridges are built, one connecting parts A and C, and the other connecting parts B and C? If your answer is positive, where should such a walk start and where should it end?



6. The figure above gives a plan of a small gallery. Is it possible to start in one room of the gallery, walk through each door of the gallery exactly once, and end at the starting point? Is it possible to start in one room, walk through each door of the gallery exactly once, and end in another room? (Here, the outdoors counts as a room, too.)



7. The figure above shows a map of a small town (streets and intersections). Is it possible to paint centerlines in all the streets without traversing any street segment more than once?

6 Trees

6.1 Trees and Their Properties

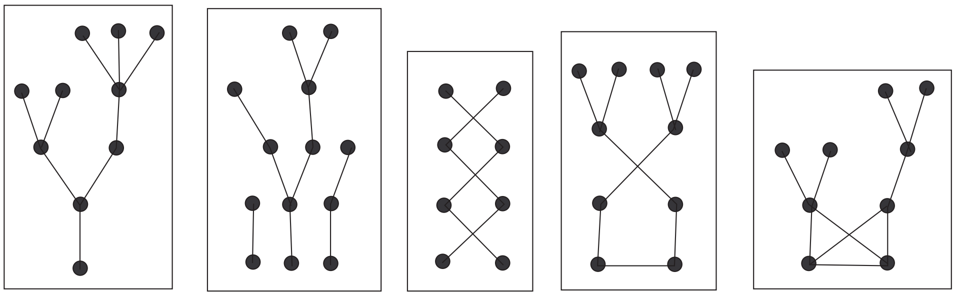
Definition 6.1 A graph without cycles is called **acyclic** or a **forest**. A **tree** is a connected acyclic graph, that is, a connected forest.

Note: Since loops and multiple edges crate cycles of length 1 and 2, respectively, a forest is necessarily a simple graph.

Example 6.2 Which of the graphs above are forests? Which are trees?
Solution: The first four graphs have no cycles, so they are forests. The first and fourth forests are also connected, so they are trees.

Trees have many nice properties. We list some of these below, some without proof.

Theorem 6.3 *Let G be a graph. Then G is a tree if and only if for any two vertices $u, v \in V(G)$, there exists a unique (u, v)-path in G.*



Theorem 6.4 *Every tree with at least 2 vertices has at least 2 vertices of degree 1 (called leaves).*

Theorem 6.5 *Any tree with n vertices has exactly $n-1$ edges.*

PROOF. By induction on the number of vertices.

Let G be a tree with 1 vertex. Then $G \cong K_1$, so it has no edges. Hence the basis of induction holds.

Assume that every tree with k vertices, for some $k \geq 1$, has exactly $k-1$ edges (induction hypothesis). Take a tree G with $k+1$ vertices. By the previous theorem (since $k+1 \geq 2$), G has a leaf, say x. Deleting x and the unique edge of G incident with x from G results in a tree G' with k vertices. By the induction hypothesis, G' has exactly $k-1$ edges, and so G has exactly k edges (one less than the number of vertices). Hence the induction step holds as well.

We conclude that, by induction, every tree with n vertices has exactly $n-1$ edges. □

6.2 Trees as Models

Like general graphs, trees have a variety of applications in many areas of science, engineering, and the social sciences. We mention just a few.

- **Social sciences:** e.g. a family tree, $G = (V, E)$ where

  $V = \{\text{people in a family}\}$

  $uv \in E \quad \Leftrightarrow \quad$ persons u and v are parent and child

- **Chemistry:** e.g. a molecular model (of a saturated hydrocarbon), $G = (V, E)$ where

  $V = \{\text{carbon and hydrogen atoms in the molecule}\}$

  $E = \{\text{chemical bonds}\}$

- **Social sciences:** e.g. a hierarchical structure of an organization, $G = (V, E)$ where

  $V = \{\text{members of the organization}\}$

  $uv \in E \quad \Leftrightarrow \quad$ persons u and v are a member and his/her immediate superior

- **Biology:** e.g. an evolutionary tree, $G = (V, E)$ where

  $V = \{\text{living species and extinct ancestor species of interest}\}$

  $uv \in E \quad \Leftrightarrow \quad$ species u and v are a species and its immediate ancestor of interest

- **Computer science:** e.g. a directory and file system, $G = (V, E)$ where

  $V = \{\text{directories and files}\}$

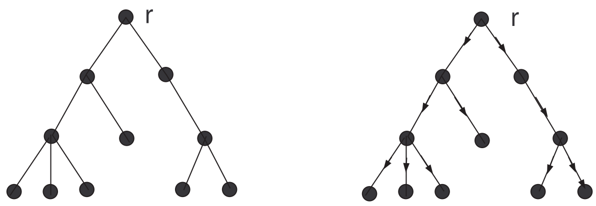
  $uv \in E \quad \Leftrightarrow \quad u$ and v are a directory and immediate subdirectory, or a directory and a file contained in it

- **Computer science:** e.g.
  - tree-connected networks of parallel processors
  - decision trees (algorithms)
  - search trees
  - prefix codes

And many many more...

6.3 Rooted Trees

Definition 6.6 A tree with a vertex designated as the *root* is called a **rooted tree**. A rooted tree can be regarded as a directed graph with all edges directed *away from* the root.



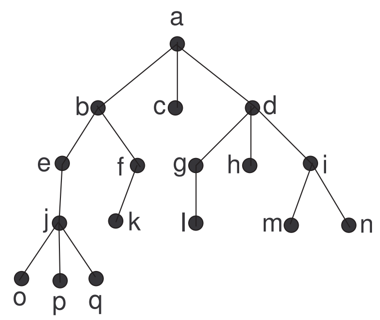
The figure above shows a tree with root r and the corresponding directed graph.

Example 6.7 Which of the trees in Section 6.2 can be considered rooted trees?

Answer: Family trees and evolutionary trees (in both cases, the root is the oldest common ancestor in the tree); hierarchical trees (the root is the top-most boss); directory trees (the root is the all-inclusive directory).

Definition 6.8 Terminology for rooted trees. Let $T = (V, E)$ be a rooted tree with root r and $u, v \in V$.

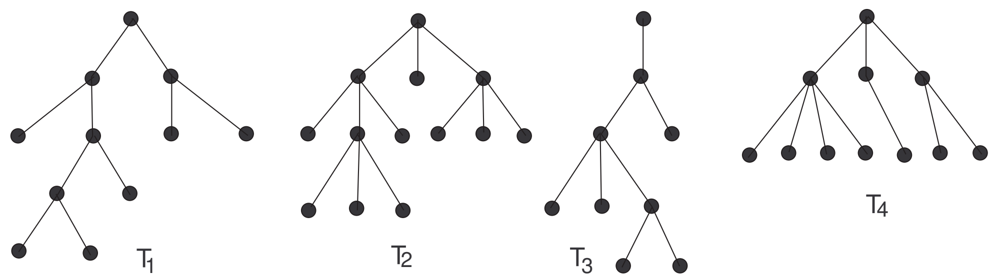
- If u lies on the unique (v, r)-path, then u is called an **ancestor** of v, and v is called a **descendant** of u.
- If u lies on the unique (v, r)-path and $u \sim v$, then u is called the **parent** of v, and v is called a **child** of u.
- If u and v have the same parent, then they are called **siblings**.
- If vertex u has a child, then it is called an **internal vertex**; if it has no children, then it is called a **leaf**.
- The subgraph of T consisting of a vertex u, all its descendants, and all incident edges is called the **subtree of T rooted at u**.



Example 6.9 In the above rooted tree with vertex a as the root, vertices b, c, and d are its children. All vertices other than a are descendants of a. Vertex j has ancestors a, b, and e. The descendants of vertex d are g, h, i, l, m, n. Vertices g, h, and i are siblings; they have the same parent d. Vertices a, b, d, e, f, g, i, j are internal vertices, while $c, h, k, l, m, n, o, p, q$ are leaves. The subtree rooted at vertex e contains vertices e, j, o, p, q and the edges that link them.

Definition 6.10 A rooted tree is called

- an m**-ary tree** if every internal vertex has at most m children; and
- a **full** m**-ary tree** if every internal vertex has exactly m children.
- A **binary tree** is a 2-ary tree.
- A **ternary tree** is a 3-ary tree.



Example 6.11 For each of the trees above, determine the smallest m so that the tree is m-ary. For this m, which of these trees are full m-ary trees?

Solution: Tree T_1 is a full binary tree, and T_2 is a full ternary tree. Tree T_3 is also ternary, but not full, while T_4 is 4-ary but not full.

Theorem 6.12 *Let T be a full m-ary tree with n vertices, of which i are internal vertices. Then*

$$n = mi + 1.$$

PROOF. We count the number of all vertices that are children in two ways. First, every vertex except the root is a child of some vertex, hence the number of children is $n-1$. On the other hand, each internal vertex has exactly m children, and each vertex is a child of exactly one parent. Hence the total number of children is mi. We conclude that $n-1=mi$, and hence $n=mi+1$. □

Example 6.13 Let T be a full m-ary tree.

1. If T has n vertices, how many of these are
   (a) internal vertices?
   (b) leaves?

   Solution: Let i and ℓ be the numbers of internal vertices and leaves, respectively. By Theorem 6.12, we have $n=mi+1$, so
   $$i=\frac{n-1}{m}.$$
   Since each vertex is either an internal vertex or a leaf, we have $n=i+\ell$. Therefore
   $$\ell=n-i=n-\frac{n-1}{m}=\frac{n(m-1)+1}{m}.$$

2. If T has i internal vertices, what is the number of
   (a) all vertices?
   (b) leaves?

   Solution: By Theorem 6.12, we have that the total number of vertices is
   $$n=mi+1.$$
   Hence the number of leaves is
   $$\ell=n-i=mi+1-i=(m-1)i+1.$$

3. If T has ℓ leaves, what is the number of
   (a) internal vertices?
   (b) all vertices?

   Solution: From the above, $\ell=(m-1)i+1$. Hence
   $$i=\frac{\ell-1}{m-1}$$
   and
   $$n=\ell+i=\ell+\frac{\ell-1}{m-1}=\frac{\ell m-1}{m-1}.$$

Example 6.14 A chain letter starts when a person sends a letter to 7 others. Each person receiving the letter either sends it to 7 others who have never received the letter before, or does not send it to anyone. Suppose 5989 people do not send out the letter before the chain is broken (and no one received more than one letter). How many received the letter? How many people sent out letters?

Solution: We have a full m-ary tree with $m = 7$. Let n be the total number of vertices, i the number of internal vertices, and ℓ the number of leaves. The number of leaves (people who did not send out the letter) is given as $\ell = 5989$. We must determine the number of people who received the letter, that is, $n - 1$, and the number of people who sent out the letter, that is, i. By the previous exercise, we have

$$i = \frac{\ell - 1}{m - 1} = \frac{5989 - 1}{7 - 1} = 998$$

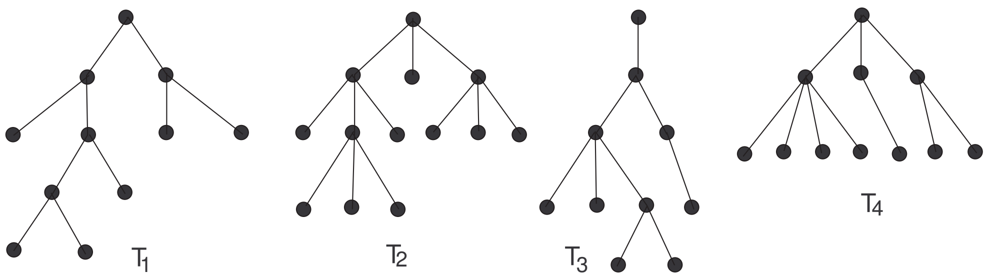
and

$$n - 1 = \ell + i - 1 = 5989 + 998 - 1 = 6986.$$

Hence 6986 people received the letter, while only 998 people sent it out.

Definition 6.15 Let T be a rooted tree with root r and vertex u.

- The length of the unique path in T from r to u is called the **level** of vertex u.
- The maximum level of any vertex in T is called the **height** of T.
- If T is of height h, then T is called **balanced** if every leaf is at level h or $h - 1$.



Example 6.16 Determine the height of each rooted tree above. Which of these trees are balanced?

Solution: Tree T_1 is of height 4; it is not balanced because we have leaves at level 2 as well. Tree T_2 is of height 3; it is not balanced because we have leaves at level 1 as well. Tree T_3 is balanced of height 4, and T_4 is balanced of height 2.

Theorem 6.17 *Let T be an m-ary tree of height h with ℓ leaves. Then:*

1. $\ell \leq m^h$.
2. *If T is a full balanced m-ary tree, then* $\ell \geq m^{h-1} + 1$.

PROOF.

1. By induction on h.

   Let T be an m-ary tree of height 0 with ℓ leaves. Then T consists of a single vertex, which is both the root and a leaf. We have $\ell = 1$ and $m^h = 0$. Hence $\ell \leq m^h$ as claimed.

   Now suppose that for some $h \geq 0$, every m-ary tree of height h with ℓ leaves satisfies $\ell \leq m^h$. Let T be an m-ary tree of height $h+1$ with root r. Now r has subtrees $T_1, \ldots, T_k$ rooted at the children of vertex r, and $k \leq m$. Each of $T_1, \ldots, T_k$ is an m-ary tree of height at most h, so by the induction hypothesis, each has at most m^h leaves. Thus T has at most $k \cdot m^h \leq m^{h+1}$ leaves.

   The result follows by induction.

2. Now assume T is a full balanced m-ary tree of height h. Since vertices at levels $0, 1, \ldots, h-2$ are all internal, there must be exactly m^i vertices at level i for each $i = 0, 1, \ldots, h-1$. Each vertex at level $h-1$ is either a leaf, or a parent of m leaves. In addition, there is at least one vertex at level h. Hence the number of leaves is at least $m^{h-1} + 1$.

□

Example 6.18 1. A full m-ary tree T has 76 leaves and height 3. What are the possible values of m?

Solution: As usual, let n, i, and ℓ denote the total number of vertices, the number of internal vertices, and the number of leaves, respectively. From Example 6.13 we have $\ell = (m-1)i + 1$. Hence

$$75 = \ell - 1 = (m-1)i.$$

Since $75 = 3 \cdot 5^2$, there are only three ways to write 75 as a product of two positive integers, namely, $75 = 1 \cdot 75 = 3 \cdot 25 = 5 \cdot 15$. Since T has height 3, the number of internal vertices is at least 3. In addition, by Theorem 6.17, $\ell \leq m^h$, that is, $76 \leq m^3$, whence $m > 4$. Therefore

$$(i, m-1) \in \{(3, 25), (5, 15), (15, 5)\}.$$

A tree with $i = 3$ and $m = 26$ exists: it has 26 vertices on level 1, of which 25 are leaves; 26 vertices on level 2, of which 25 are leaves, and 26 vertices (all leaves) on level 3.

A tree with $i = 5$ and $m = 16$ exists: for example, it can have 16 vertices on level 1, of which 13 are leaves; 48 vertices on level 2, of which 47 are leaves, and 16 vertices (all leaves) on level 3.

A tree with $i = 15$ and $m = 6$ also exists: for example, it can have 6 vertices on level 1, all internal vertices; 36 vertices on level 2, of which 28 are leaves, and 48 vertices (all leaves) on level 3.

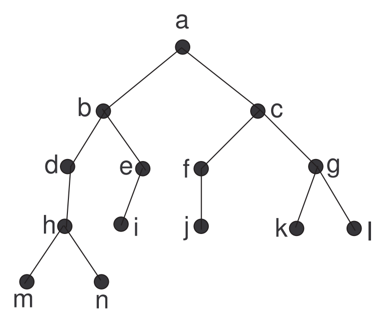
2. What if T is also balanced?

   Solution: By Theorem 6.17, we need $76 = \ell > m^{h-1} = m^2$. Hence $m \leq 8$. Thus $m = 6$ from (a). As seen above, a balanced full 6-ary tree with 15 internal vertices and 76 leaves indeed exists.

Definition 6.19 An **ordered** rooted tree is a rooted tree where the children of each internal vertex are *ordered.*

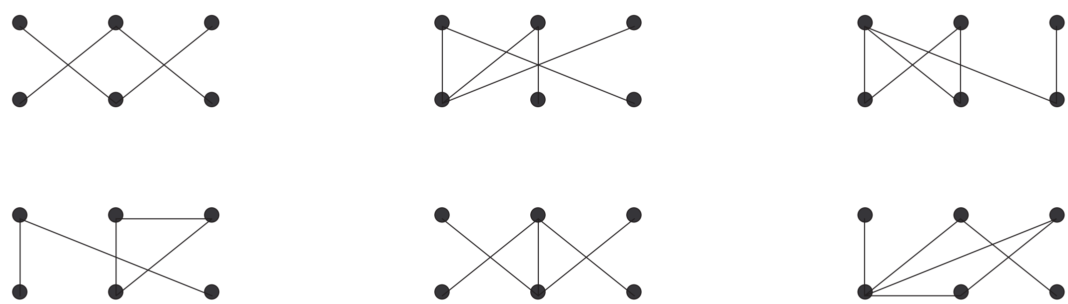
Let T be an ordered binary tree. Then:

- The first child of an internal vertex u is called the **left child** of u, and the second is called the **right child** of u.
- The subtree of T rooted at the left child of vertex u is called the **left subtree** of u, and the subtree of T rooted at the right child of vertex u is called the **right subtree** of u.

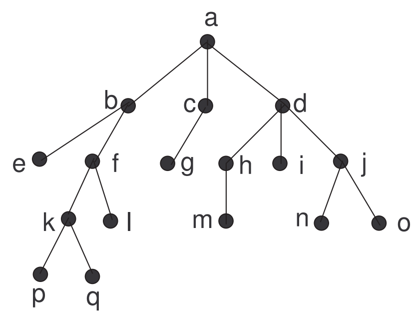


Example 6.20 Consider the ordered tree above, with the drawing determining the order of children. Vertices b and c are the left and right child of the root a, respectively. The subtrees rooted at d and e are the left and right subtree of b, respectively.

6.4 Exercises



1. Which of the six graphs above are trees? Which are disconnected forests?
2. Find all pairwise non-isomorphic trees with (a) 4 vertices, (b) 6 vertices.
3. Which complete bipartite graphs $K_{m,n}$ are trees?



4. How many vertices does a tree with 2015 edges have?

5. Consider the rooted tree T above.

   (a) Describe the tree T using the following terms: root, internal vertex, leaf; child, parent, sibling; ancestor, descendant; subtree.

   (b) Determine the smallest m so that T is an m-ary tree. Is T a full m-ary tree for this m? If not, what is the minimum number of vertices we need to adjoin, and how, to create a full m-ary tree from T?

   (c) Determine the level of each vertex in T, and the height of T. Is T balanced? If not, what is the minimum number of vertices we need to adjoin, and how, to create a balanced tree from T?

6. For each of the rooted trees described below, determine the number of all vertices, the number of edges, the number of internal vertices, and the number of leaves.

   (a) A full 5-ary tree with 100 internal vertices.

   (b) A full binary tree with 1001 leaves.

   (c) A full ternary tree with 124 vertices.

7. A chain letter starts when a person sends a letter to 5 others. Each person receiving the letter either sends it to 5 others who have never received the letter before, or does not send it to anyone. Suppose 10,000 people send out the letter before the chain is broken (and no one received more than one letter). How many received the letter? How many people did not send out letters?

8. A chain letter starts with a person sending a letter out to 10 others. Each person is asked to send the letter to 10 others, and each letter contains a list of the previous 6 people in the chain. Unless there are fewer than 6 people on the list, each person sends \$1 to the first person on the list, deletes this persons name from the list, moves up each of the other 5 names one position, and inserts his/her name at the end of the list. If no person breaks the chain and no one receives more than one letter, how much money will a person in the chain ultimately receive?

9. Draw a full m-ary tree with 84 leaves and height 3, or else show that such a tree does not exist.

10. A full m-ary tree T has 81 leaves and height 4.

    (a) Give an upper and a lower bound for m.

    (b) What is m if T is also balanced?