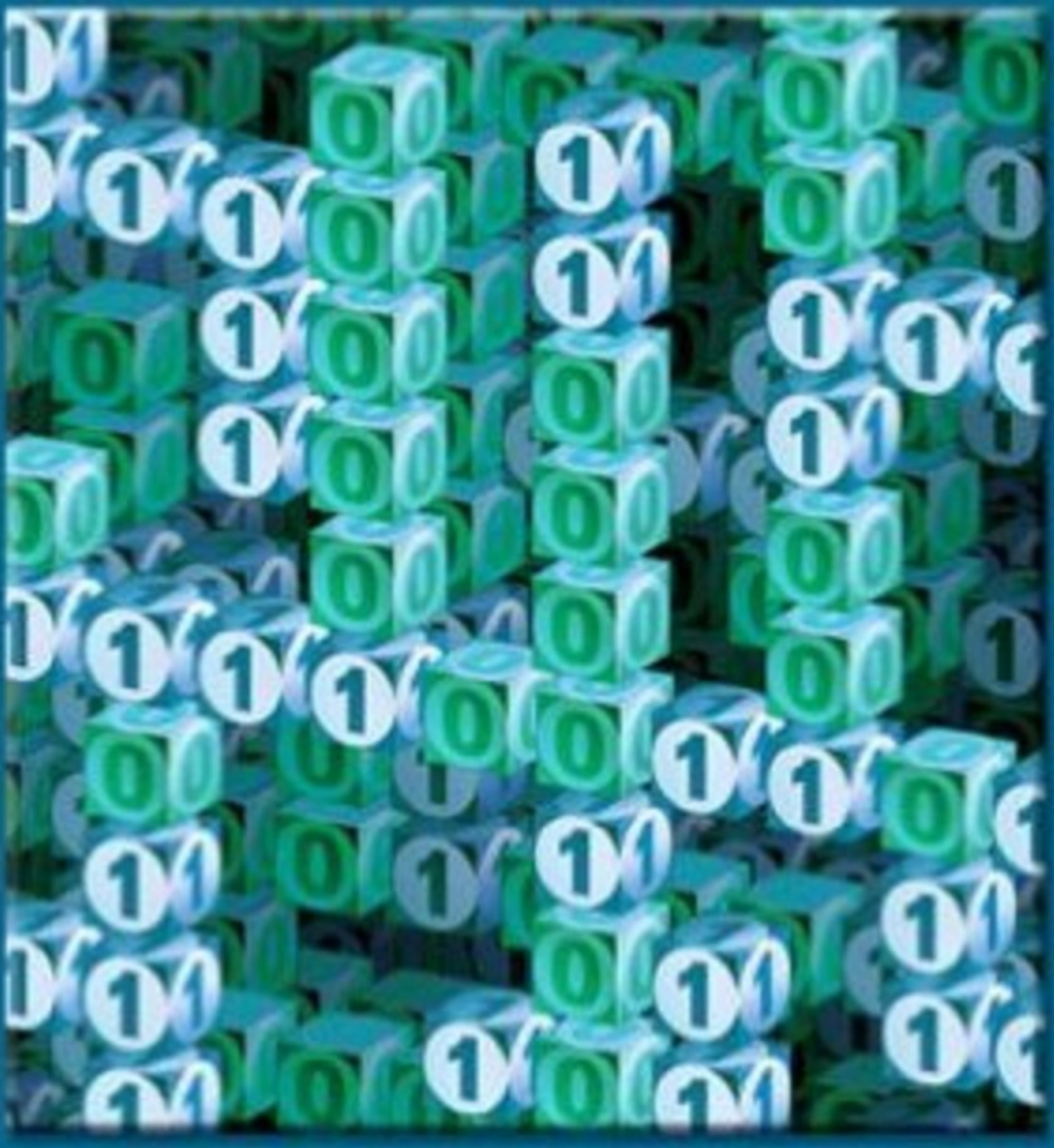
DISCRETE and COMBINATORIAL MATHEMATICS

An Applied Introduction



Ralph P. Grimaldi

Fifth Edition

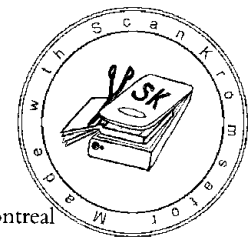
DISCRETE AND COMBINATORIAL MATHEMATICS

An Applied Introduction

FIFTH EDITION

RALPH P. GRIMALDI
Rose-Hulman Institute of Technology

Boston San Francisco New York
London Toronto Sydney Tokyo Singapore Madrid
Mexico City Munich Paris Cape Town Hong Kong Montreal



NOTATION

LOGIC	p, q	statements (or propositions)
	$\neg p$	the negation of (statement) p: *not* p
	$p \wedge q$	the conjunction of p, q: p *and* q
	$p \vee q$	the disjunction of p, q: p *or* q
	$p \rightarrow q$	the implication of q by p: p *implies* q
	$p \leftrightarrow q$	the biconditional of p and q: p *if and only if* q
	iff	if and only if
	$p \Rightarrow q$	logical implication: p *logically implies* q
	$p \Longleftrightarrow q$	logical equivalence: p is *logically equivalent* to q
	T_0	tautology
	F_0	contradiction
	$\forall x$	For *all* x (the universal quantifier)
	$\exists x$	For *some* x (the existential quantifier)
SET THEORY	$x \in A$	element x is a member of set A
	$x \notin A$	element x is not a member of set A
	$\mathcal{U}$	the universal set
	$A \subseteq B$, $B \supseteq A$	A is a subset of B
	$A \subset B$, $B \supset A$	A is a proper subset of B
	$A \nsubseteq B$	A is not a subset of B
	$A \not\subset B$	A is not a proper subset of B
	$\lvert A \rvert$	the cardinality, or size, of set A — that is, the number of elements in A
	$\emptyset = \{\,\}$	the empty, or null, set
	$\mathcal{P}(A)$	the power set of A — that is, the collection of all subsets of A
	$A \cap B$	the intersection of sets A, B: $\{x \mid x \in A \text{ and } x \in B\}$
	$A \cup B$	the union of sets A, B: $\{x \mid x \in A \text{ or } x \in B\}$
	$A \triangle B$	the symmetric difference of sets A, B: $\{x \mid x \in A \text{ or } x \in B, \text{ but } x \notin A \cap B\}$
	$\overline{A}$	the complement of set A: $\{x \mid x \in \mathcal{U} \text{ and } x \notin A\}$
	$A - B$	the (relative) complement of set B in set A: $\{x \mid x \in A \text{ and } x \notin B\}$
	$\bigcup_{i \in I} A_i$	$\{x \mid x \in A_i, \text{ for at least one } i \in I\}$, where I is an index set
	$\bigcap_{i \in I} A_i$	$\{x \mid x \in A_i, \text{ for every } i \in I\}$, where I is an index set
PROBABILITY	S	the sample space for an experiment $\mathcal{E}$
	$A \subseteq S$	A is an event
	$Pr(A)$	the probability of event A
	$Pr(A \mid B)$	the probability of A given B; conditional probability
	X	random variable
	$E(X)$	the expected value of X, a random variable
	$\text{Var}(X) = \sigma_X^2$	the variance of X, a random variable
	σ_X	the standard deviation of X, a random variable
NUMBERS	$a \mid b$	a divides b, for $a, b \in \mathbf{Z}, a \neq 0$
	$a \nmid b$	a does not divide b, for $a, b \in \mathbf{Z}, a \neq 0$
	$\gcd(a, b)$	the greatest common divisor of the integers a, b
	$\text{lcm}(a, b)$	the least common multiple of the integers a, b
	$\phi(n)$	Euler's phi function for $n \in \mathbf{Z}^+$
	$\lfloor x \rfloor$	the greatest integer less than or equal to the real number x: the greatest integer in x: the *floor* of x

NOTATION

	$\lceil x \rceil$	the smallest integer greater than or equal to the real number x: the *ceiling* of x
	$a \equiv b \pmod{n}$	a is congruent to b modulo n
RELATIONS	$A \times B$	the Cartesian, or cross, product of sets A, B: $\{(a, b) \mid a \in A, b \in B\}$
	$\mathcal{R} \subseteq A \times B$	$\mathcal{R}$ is a relation from A to B
	$a \mathcal{R} b$; $(a, b) \in \mathcal{R}$	a is related to b
	$a \not\mathcal{R} b$; $(a, b) \notin \mathcal{R}$	a is not related to b
	$\mathcal{R}^c$	the converse of relation $\mathcal{R}$: $(a, b) \in \mathcal{R}$ iff $(b, a) \in \mathcal{R}^c$
	$\mathcal{R} \circ \mathcal{S}$	the composite relation for $\mathcal{R} \subseteq A \times B$, $\mathcal{S} \subseteq B \times C$: $(a, c) \in \mathcal{R} \circ \mathcal{S}$ if $(a, b) \in \mathcal{R}$, $(b, c) \in \mathcal{S}$ for some $b \in B$
	$\text{lub}\{a, b\}$	the least upper bound of a and b
	$\text{glb}\{a, b\}$	the greatest lower bound of a and b
	$[a]$	the equivalence class of element a (relative to an equivalence relation $\mathcal{R}$ on a set A): $\{x \in A \mid x \mathcal{R} a\}$
FUNCTIONS	$f: A \to B$	f is a function from A to B
	$f(A_1)$	for $f: A \to B$ and $A_1 \subseteq A$, $f(A_1)$ is the image of A_1 under f — that is, $\{f(a) \mid a \in A_1\}$
	$f(A)$	for $f: A \to B$, $f(A)$ is the range of f
	$f: A \times A \to B$	f is a binary operation on A
	$f: A \times A \to B\ (\subseteq A)$	f is a closed binary operation on A
	$1_A: A \to A$	the identity function on A: $1_A(a) = a$ for each $a \in A$
	$f\|_{A_1}$	the restriction of $f: A \to B$ to $A_1 \subseteq A$
	$g \circ f$	the composite function for $f: A \to B$, $g: B \to C$: $(g \circ f)a = g(f(a))$, for $a \in A$
	f^{-1}	the inverse of function f
	$f^{-1}(B_1)$	the preimage of $B_1 \subseteq B$ for $f: A \to B$
	$f \in O(g)$	f is "big Oh" of g; f is of order g
THE ALGEBRA OF STRINGS	Σ	a finite set of symbols called an alphabet
	λ	the empty string
	$\|\|x\|\|$	the length of string x
	Σ^n	$\{x_1 x_2 \cdots x_n \mid x_i \in \Sigma\}$, $n \in \mathbf{Z}^+$
	Σ^0	$\{\lambda\}$
	Σ^+	$\bigcup_{n \in \mathbf{Z}^+} \Sigma^n$: the set of all strings of positive length
	Σ^*	$\bigcup_{n \geq 0} \Sigma^n$: the set of all finite strings
	$A \subseteq \Sigma^*$	A is a language
	AB	the concatenation of languages A, $B \subseteq \Sigma^*$: $\{ab \mid a \in A, b \in B\}$
	A^n	$\{a_1 a_2 \cdots a_n \mid a_i \in A \subseteq \Sigma^*\}$, $n \in \mathbf{Z}^+$
	A^0	$\{\lambda\}$
	A^+	$\bigcup_{n \in \mathbf{Z}^+} A^n$
	A^*	$\bigcup_{n \geq 0} A^n$: the Kleene closure of language A
	$M = (S, \mathcal{S}, \mathcal{O}, \nu, \omega)$	a finite state machine M with internal states S, input alphabet $\mathcal{S}$, output alphabet $\mathcal{O}$, next state function $\nu: S \times \mathcal{S} \to S$ and output function $\omega: S \times \mathcal{S} \to \mathcal{O}$

PART

3

GRAPH THEORY AND APPLICATIONS

11

An Introduction to Graph Theory

With this chapter we start to develop another major topic of this text. Unlike other areas in mathematics, the theory of graphs has a definite starting place, a paper published in 1736 by the Swiss mathematician Leonhard Euler (1707–1783). The main idea behind this work grew out of a now-popular problem known as the seven bridges of Königsberg. We shall examine the solution of this problem, from which Euler developed some of the fundamental concepts for the theory of graphs.

Unlike the continuous graphs of early algebra courses, the graphs we examine here are finite in structure and can be used to analyze relationships and applications in many different settings. We have seen some examples of applications of graph theory in earlier chapters (3, 5–8, and 10). However, the development here is independent of these prior discussions.

11.1 Definitions and Examples

When we use a road map, we are often concerned with seeing how to get from one town to another by means of the roads indicated on the map. Consequently, we are dealing with two distinct sets of objects: towns and roads. As we have seen many times before, such sets of objects can be used to define a relation. If V denotes the set of towns and E the set of roads, we can define a relation $\mathcal{R}$ on V by $a \mathcal{R} b$ if we can travel from a to b using only the roads in E. If the roads in E that take us from a to b are all two-way roads, then we also have $b \mathcal{R} a$. Should all the roads under consideration be two-way, we have a symmetric relation.

One way to represent a relation is by listing the ordered pairs that are its elements. Here, however, it is more convenient to use a picture, as shown in Fig. 11.1. This figure demonstrates the possible ways of traveling among six towns using the eight roads indicated. It shows that there is at least one set of roads connecting any two towns (identical or distinct). This pictorial representation is a lot easier to work with than the 36 ordered pairs of the relation $\mathcal{R}$.

At the same time, Fig. 11.1 would be appropriate for representing six communication centers, with the eight "roads" interpreted as communication links. If each link provides two-way communication, we should be quite concerned about the vulnerability of center a to such hazards as equipment breakdown or enemy attack. Without center a, neither b nor c can communicate with any of d, e, or f.

From these observations we consider the following concepts.

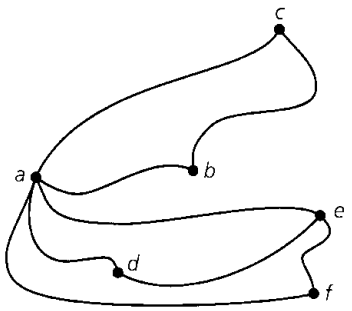


Figure 11.1

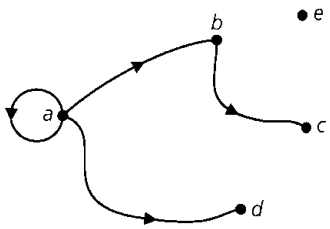


Figure 11.2

Definition 11.1 Let V be a finite nonempty set, and let $E \subseteq V \times V$. The pair (V, E) is then called a *directed graph* (on V), or *digraph*[†] (on V), where V is the set of *vertices*, or *nodes*, and E is its set of (*directed*) *edges* or *arcs*. We write $G = (V, E)$ to denote such a graph.

When there is no concern about the direction of any edge, we still write $G = (V, E)$. But now E is a set of unordered pairs of elements taken from V, and G is called an *undirected graph*.

Whether $G = (V, E)$ is directed or undirected, we often call V the *vertex set* of G and E the *edge set* of G.

Figure 11.2 provides an example of a directed graph on $V = \{a, b, c, d, e\}$ with $E = \{(a, a), (a, b), (a, d), (b, c)\}$. The direction of an edge is indicated by placing a directed arrow on the edge, as shown here. For any edge, such as (b, c), we say that the edge is *incident* with the vertices b, c; b is said to be *adjacent to* c, whereas c is *adjacent from* b. In addition, vertex b is called the *origin*, or *source*, of the edge (b, c), and vertex c is the *terminus*, or *terminating vertex*. The edge (a, a) is an example of a *loop*, and the vertex e that has no incident edges is called an *isolated* vertex.

An undirected graph is shown in Fig. 11.3(a). This graph is a more compact way of describing the directed graph given in Fig. 11.3(b). In an undirected graph, there are undirected edges such as $\{a, b\}, \{b, c\}, \{a, c\}, \{c, d\}$ in Fig. 11.3(a). An edge such as $\{a, b\}$ stands for $\{(a, b), (b, a)\}$. Although $(a, b) = (b, a)$ only when $a = b$, we do have $\{a, b\} = \{b, a\}$

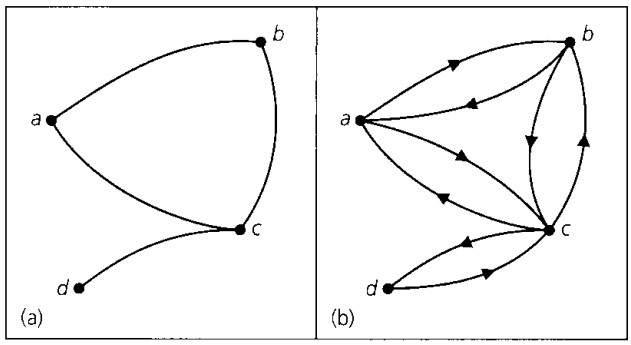


Figure 11.3

[†] Since the terminology of graph theory is not standard, the reader may find some differences between terms used here and in other texts.

for any a, b. We can write $\{a, a\}$ to denote a loop in an undirected graph, but $\{a, a\}$ is considered the same as (a, a).

In general, if a graph G is not specified as directed or undirected, it is assumed to be undirected. When it contains no loops it is called *loop-free.*

In the next two definitions we shall not concern ourselves with any loops that may be present in the undirected graph G.

Definition 11.2

Let x, y be (not necessarily distinct) vertices in an undirected graph $G = (V, E)$. An x-y *walk* in G is a (loop-free) finite alternating sequence

$$x = x_0, e_1, x_1, e_2, x_2, e_3, \ldots, e_{n-1}, x_{n-1}, e_n, x_n = y$$

of vertices and edges from G, starting at vertex x and ending at vertex y and involving the n edges $e_i = \{x_{i-1}, x_i\}$, where $1 \leq i \leq n$.

The *length* of this walk is n, the number of edges in the walk. (When $n = 0$, there are no edges, $x = y$, and the walk is called *trivial.* These walks are not considered very much in our work.)

Any x-y walk where $x = y$ (and $n > 1$) is called a *closed walk.* Otherwise the walk is called *open.*

Note that a walk may repeat both vertices and edges.

EXAMPLE 11.1

For the graph in Fig. 11.4 we find, for example, the following three open walks. We can list the edges only or the vertices only (if the other is clearly implied).

1) $\{a, b\}, \{b, d\}, \{d, c\}, \{c, e\}, \{e, d\}, \{d, b\}$: This is an a-b walk of length 6 in which we find the vertices d and b repeated, as well as the edge $\{b, d\}$ $(= \{d, b\})$.
2) $b \to c \to d \to e \to c \to f$: Here we have a b-f walk where the length is 5 and the vertex c is repeated, but no edge appears more than once.
3) $\{f, c\}, \{c, e\}, \{e, d\}, \{d, a\}$: In this case the given f-a walk has length 4 with no repetition of either vertices or edges.

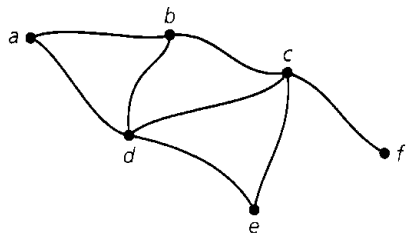


Figure 11.4

Since the graph of Fig. 11.4 is undirected, the a-b walk in part (1) is also a b-a walk (we read the edges, if necessary, as $\{b, d\}, \{d, e\}, \{e, c\}, \{c, d\}, \{d, b\}$, and $\{b, a\}$). Similar remarks hold for the walks in parts (2) and (3).

Finally, the edges $\{b, c\}$, $\{c, d\}$, and $\{d, b\}$ provide a b-b (closed) walk. These edges (ordered appropriately) also define (closed) c-c and d-d walks.

Now let us examine special types of walks.

Definition 11.3 Consider any x-y walk in an undirected graph $G = (V, E)$.

a) If no edge in the x-y walk is repeated, then the walk is called an x-y *trail*. A closed x-x trail is called a *circuit*.

b) If no vertex of the x-y walk occurs more than once, then the walk is called an x-y *path*. When $x = y$, the term *cycle* is used to describe such a closed path.

Convention: In dealing with circuits, we shall always understand the presence of at least one edge. When there is only one edge, then the circuit is a loop (and the graph is no longer loop-free). Circuits with two edges arise in multigraphs, a concept we shall define shortly.

The term *cycle* will always imply the presence of at least three distinct edges (from the graph).

EXAMPLE 11.2

a) The b-f walk in part (2) of Example 11.1 is a b-f trail, but it is not a b-f path because of the repetition of vertex c. However, the f-a walk in part (3) of that example is both an f-a trail (of length 4) and an f-a path (of length 4).

b) In Fig. 11.4, the edges $\{a, b\}$, $\{b, d\}$, $\{d, c\}$, $\{c, e\}$, $\{e, d\}$, and $\{d, a\}$ provide an a-a circuit. The vertex d is repeated, so the edges do *not* give us an a-a cycle.

c) The edges $\{a, b\}$, $\{b, c\}$, $\{c, d\}$, and $\{d, a\}$ provide an a-a cycle (of length 4) in Fig. 11.4. When ordered appropriately these same edges may also define a b-b, c-c, or d-d cycle. Each of these cycles is also a circuit.

For a directed graph we shall use the adjective *directed*, as in, for example, *directed walks, directed paths*, and *directed cycles*.

Before continuing, we summarize (in Table 11.1) for future reference the results of Definitions 11.2 and 11.3. Each occurrence of "Yes" in the first two columns here should be interpreted as "Yes, possibly." Table 11.1 reflects the fact that a path is a trail, which in turn is an open walk. Furthermore, every cycle is a circuit, and every circuit (with at least two edges) is a closed walk.

Table 11.1

Repeated Vertex (Vertices)	Repeated Edge(s)	Open	Closed	Name
Yes	Yes	Yes		Walk (open)
Yes	Yes		Yes	Walk (closed)
Yes	No	Yes		Trail
Yes	No		Yes	Circuit
No	No	Yes		Path
No	No		Yes	Cycle

Considering how many concepts we have introduced, it is time to prove a first result in this new theory.

THEOREM 11.1

Let $G = (V, E)$ be an undirected graph, with $a, b \in V, a \neq b$. If there exists a trail (in G) from a to b, then there is a path (in G) from a to b.

Proof: Since there is a trail from a to b, we select one of shortest length, say $\{a, x_1\}, \{x_1, x_2\}, \ldots, \{x_n, b\}$. If this trail is not a path, we have the situation $\{a, x_1\}, \{x_1, x_2\}, \ldots, \{x_{k-1}, x_k\}, \{x_k, x_{k+1}\}, \{x_{k+1}, x_{k+2}\}, \ldots, \{x_{m-1}, x_m\}, \{x_m, x_{m+1}\}, \ldots, \{x_n, b\}$, where $k < m$ and $x_k = x_m$, possibly with $k = 0$ and $a\ (= x_0) = x_m$, or $m = n + 1$ and $x_k = b\ (= x_{n+1})$. But then we have a contradiction because $\{a, x_1\}, \{x_1, x_2\}, \ldots, \{x_{k-1}, x_k\}, \{x_m, x_{m+1}\}, \ldots, \{x_n, b\}$ is a shorter trail from a to b.

The notion of a path is needed in the following graph property.

Definition 11.4

Let $G = (V, E)$ be an undirected graph. We call G *connected* if there is a path between any two distinct vertices of G.

Let $G = (V, E)$ be a directed graph. Its associated undirected graph is the graph obtained from G by ignoring the directions on the edges. If more than one undirected edge results for a pair of distinct vertices in G, then only one of these edges is drawn in the associated undirected graph. When this associated graph is connected, we consider G connected.

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

The graphs in Figs. 11.1, 11.3, and 11.4 are connected. In Fig. 11.2 the graph is not connected because, for example, there is no path from a to e.

EXAMPLE 11.3

In Fig. 11.5 we have an undirected graph on $V = \{a, b, c, d, e, f, g\}$. This graph is not connected because, for example, there is no path from a to e. However, the graph is composed of pieces (with vertex sets $V_1 = \{a, b, c, d\}$, $V_2 = \{e, f, g\}$, and edge sets $E_1 = \{\{a, b\}, \{a, c\}, \{a, d\}, \{b, d\}\}$, $E_2 = \{\{e, f\}, \{f, g\}\}$) that are themselves connected, and these pieces are called the (*connected*) *components* of the graph. Hence an undirected graph $G = (V, E)$ is disconnected if and only if V can be partitioned into at least two subsets V_1, V_2 such that there is no edge in E of the form $\{x, y\}$, where $x \in V_1$ and $y \in V_2$. A graph is connected if and only if it has only one component.

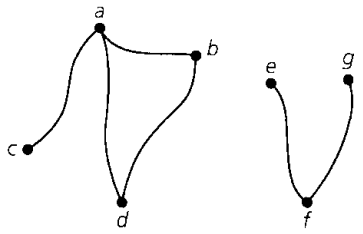


Figure 11.5

Definition 11.5

For any graph $G = (V, E)$, the number of components of G is denoted by $\kappa(G)$.

EXAMPLE 11.4

For the graphs in Figs. 11.1, 11.3, and 11.4, $\kappa(G) = 1$ because these graphs are connected; $\kappa(G) = 2$ for the graphs in Figs. 11.2 and 11.5.

Before closing this first section, we extend our concept of a gràph. Thus far we have allowed at most one edge between two vertices; we now consider an extension.

Definition 11.6 Let V be a finite nonempty set. We say that the pair (V, E) determines a *multigraph* G with vertex set V and edge set E[†] if, for some $x, y \in V$, there are two or more edges in E of the form (a) (x, y) (for a directed multigraph), or (b) $\{x, y\}$ (for an undirected multigraph). In either case, we write $G = (V, E)$ to designate the multigraph, just as we did for graphs.

Figure 11.6 shows an example of a directed multigraph. There are three edges from a to b, so we say that the edge (a, b) has *multiplicity* 3. The edges (b, c) and (d, e) both have multiplicity 2. Also, the edge (e, d) and either one of the edges (d, e) form a (directed) circuit of length 2 in the multigraph.

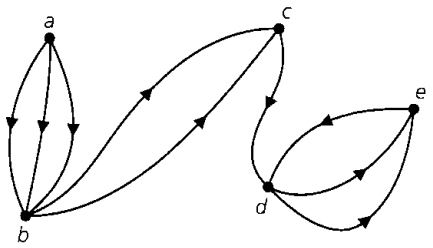


Figure 11.6

We shall need the idea of a multigraph later in the chapter when we solve the problem of the seven bridges of Königsberg. (*Note*: Whenever we are dealing with a multigraph G, we shall state explicitly that G is a multigraph.)

EXERCISES 11.1

1. List three situations, different from those in this section, where a graph could prove useful.

2. For the graph in Fig. 11.7, determine (a) a walk from b to d that is not a trail; (b) a b-d trail that is not a path; (c) a path from b to d; (d) a closed walk from b to b that is not a circuit; (e) a circuit from b to b that is not a cycle; and (f) a cycle from b to b.

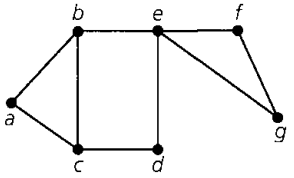


Figure 11.7

3. For the graph in Fig. 11.7, how many paths are there from b to f?

4. For $n \geq 2$, let $G = (V, E)$ be the loop-free undirected graph, where V is the set of binary n-tuples (of 0's and 1's) and $E = \{\{v, w\} | v, w \in V$ and v, w differ in (exactly) two positions$\}$. Find $\kappa(G)$.

5. Let $G = (V, E)$ be the undirected graph in Fig. 11.8. How many paths are there in G from a to h? How many of these paths have length 5?

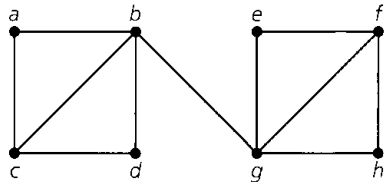


Figure 11.8

6. If a, b are distinct vertices in a connected undirected graph G, the *distance* from a to b is defined to be the length of a shortest path from a to b (when $a = b$ the *distance* is defined to be

[†]We now allow a set to have repeated elements in order to account for multiple edges. We realize that this is a change from the way we dealt with sets in Chapter 3. To overcome this the term *multiset* is often used to describe E in this case.

0). For the graph in Fig. 11.9, find the distances from d to (each of) the other vertices in G.

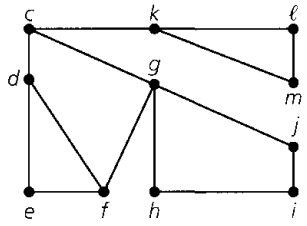


Figure 11.9

7. Seven towns a, b, c, d, e, f, and g are connected by a system of highways as follows: (1) I-22 goes from a to c, passing through b; (2) I-33 goes from c to d and then passes through b as it continues to f; (3) I-44 goes from d through e to a; (4) I-55 goes from f to b, passing through g; and (5) I-66 goes from g to d.

a) Using vertices for towns and directed edges for segments of highways between towns, draw a directed graph that models this situation.

b) List the paths from g to a.

c) What is the smallest number of highway segments that would have to be closed down in order for travel from b to d to be disrupted?

d) Is it possible to leave town c and return there, visiting each of the other towns only once?

e) What is the answer to part (d) if we are not required to return to c?

f) Is it possible to start at some town and drive over each of these highways exactly once? (You are allowed to visit a town more than once, and you need not return to the town from which you started.)

8. Figure 11.10 shows an undirected graph representing a section of a department store. The vertices indicate where cashiers are located; the edges denote unblocked aisles between cashiers. The department store wants to set up a security system where (plainclothes) guards are placed at certain cashier locations so that each cashier either has a guard at his or her location or is only one aisle away from a cashier who has a guard. What is the smallest number of guards needed?

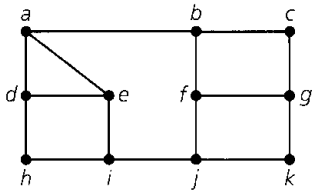


Figure 11.10

9. Let $G = (V, E)$ be a loop-free connected undirected graph, and let $\{a, b\}$ be an edge of G. Prove that $\{a, b\}$ is part of a cycle if and only if its removal (the vertices a and b are left) does not disconnect G.

10. Give an example of a connected graph G where removing any edge of G results in a disconnected graph.

11. Let G be a graph that satisfies the condition in Exercise 10. (a) Must G be loop-free? (b) Could G be a multigraph? (c) If G has n vertices, can we determine how many edges it has?

12. a) If $G = (V, E)$ is an undirected graph with $|V| = v$, $|E| = e$, and no loops, prove that $2e \leq v^2 - v$.

b) State the corresponding inequality for the case when G is directed.

13. Let $G = (V, E)$ be an undirected graph. Define a relation $\mathcal{R}$ on V by $a \mathcal{R} b$ if $a = b$ or if there is a path in G from a to b. Prove that $\mathcal{R}$ is an equivalence relation. Describe the partition of V induced by $\mathcal{R}$.

14. a) Consider the three connected undirected graphs in Fig. 11.11. The graph in part (a) of the figure consists of a cycle (on the vertices u_1, u_2, u_3) and a vertex u_4 with edges (spokes) drawn from u_4 to the other three vertices. This graph is called the *wheel with three spokes* and is denoted by W_3. In part (b) of the figure we find the graph

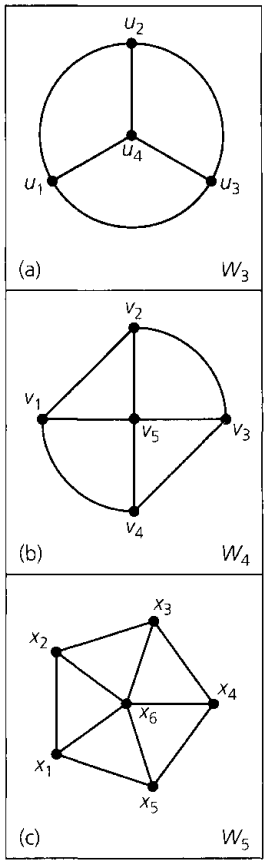


Figure 11.11

W_4 — the wheel with four spokes. The wheel W_5 with five spokes appears in Fig. 11.11(c). Determine how many cycles of length 4 there are in each of these graphs.

b) In general, if $n \in \mathbf{Z}^+$ and $n \geq 3$, then the *wheel with n spokes* is the graph made up of a cycle of length n together with an additional vertex that is adjacent to the n vertices of the cycle. The graph is denoted by W_n. (i) How many cycles of length 4 are there in W_n? (ii) How many cycles in W_n have length n?

15. For the undirected graph in Fig. 11.12, find and solve a recurrence relation for the number of closed v-v walks of length $n \geq 1$, if we allow such a walk, in this case, to contain or consist of one or more loops.

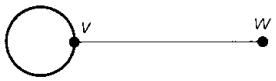


Figure 11.12

16. *Unit-Interval Graphs.* For $n \geq 1$, we start with n closed intervals of unit length and draw the corresponding unit-interval graph on n vertices, as shown in Fig. 11.13. In part (a) of the figure we have one unit interval. This corresponds to the single vertex u; both the interval and the unit-interval graph can be represented by the binary sequence 01. In parts (b), (c) of the figure we have the two unit-interval graphs determined by two unit intervals. When two unit intervals overlap [as in part (c)] an edge is drawn in the unit-interval graph joining the vertices corresponding to these unit intervals. Hence the unit-interval graph in part (b) consists of the two isolated vertices v_1, v_2 that correspond with the nonoverlapping unit intervals. In part (c) the unit intervals overlap so the corresponding unit-interval graph consists of a single edge joining the vertices v_1, v_2 (that correspond to the given unit intervals). A closer look at the unit intervals in part (c) reveals how we can represent the positioning of these intervals and the corresponding unit-interval graph by the binary sequence 0011. In parts (d)–(f) of the figure we have three of the unit-interval graphs for three unit intervals — together with their corresponding binary sequences.

a) How many other unit-interval graphs are there for three unit intervals? What are the corresponding binary sequences for these graphs?

b) How many unit-interval graphs are there for four unit intervals?

c) For $n \geq 1$, how many unit-interval graphs are there for n unit intervals?

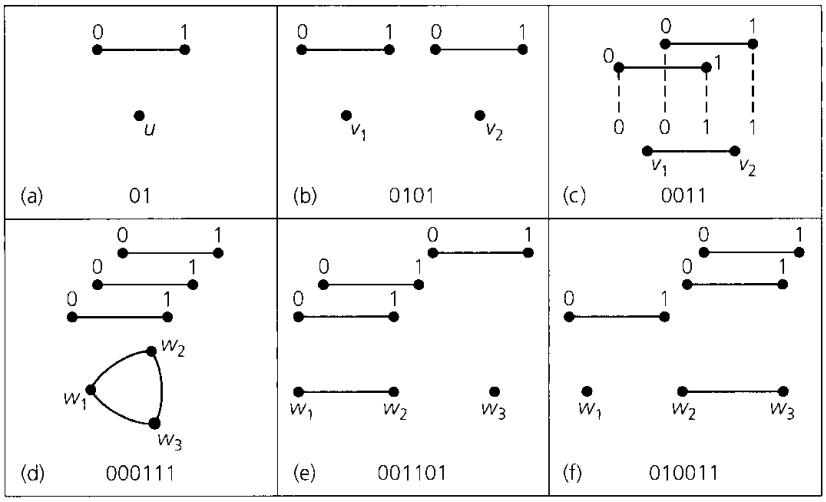


Figure 11.13

11.2 Subgraphs, Complements, and Graph Isomorphism

In this section we shall focus on the following two ideas:

a) What types of substructures are present in a graph?

b) Is it possible to draw two graphs that appear distinct but have the same underlying structure?

To answer the question in part (a) we introduce the following definition.

Definition 11.7 If $G = (V, E)$ is a graph (directed or undirected), then $G_1 = (V_1, E_1)$ is called a *subgraph* of G if $\emptyset \neq V_1 \subseteq V$ and $E_1 \subseteq E$, where each edge in E_1 is incident with vertices in V_1.

Figure 11.14(a) provides us with an undirected graph G and two of its subgraphs, G_1 and G_2. The vertices a, b are isolated in subgraph G_1. Part (b) of the figure provides a directed example. Here vertex w is isolated in the subgraph G'.

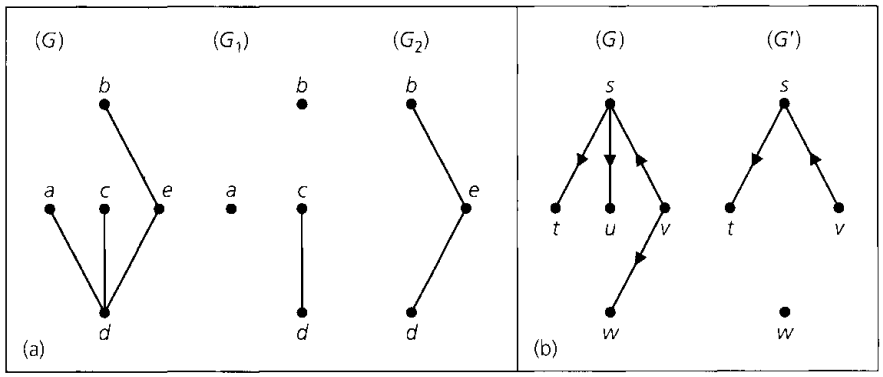


Figure 11.14

Certain special types of subgraphs arise as follows:

Definition 11.8 Given a (directed or undirected) graph $G = (V, E)$, let $G_1 = (V_1, E_1)$ be a subgraph of G. If $V_1 = V$, then G_1 is called a *spanning subgraph* of G.

In part (a) of Fig. 11.14 neither G_1 nor G_2 is a spanning subgraph of G. The subgraphs G_3 and G_4 — shown in part (a) of Fig. 11.15 — are both spanning subgraphs of G. The directed graph G' in part (b) of Fig. 11.14 is a subgraph, but *not* a spanning subgraph, of the directed graph G given in that part of the figure. In part (b) of Fig. 11.15 the directed graphs G'' and G''' are two of the $2^4 = 16$ possible spanning subgraphs.

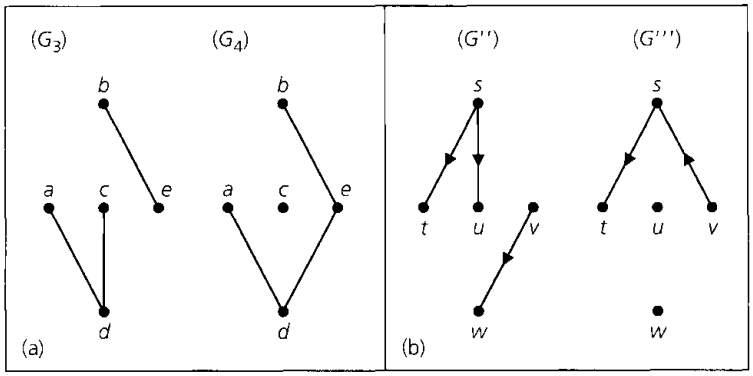


Figure 11.15

Definition 11.9

Let $G = (V, E)$ be a graph (directed or undirected). If $\emptyset \neq U \subseteq V$, the *subgraph of G induced by U* is the subgraph whose vertex set is U and which contains all edges (from G) of either the form (a) (x, y), for $x, y \in U$ (when G is directed), or (b) $\{x, y\}$, for $x, y \in U$ (when G is undirected). We denote this subgraph by $\langle U \rangle$.

A subgraph G' of a graph $G = (V, E)$ is called an *induced subgraph* if there exists $\emptyset \neq U \subseteq V$, where $G' = \langle U \rangle$.

For the subgraphs in Fig. 11.14(a), we find that G_2 is an induced subgraph of G but the subgraph G_1 is not an induced subgraph because edge $\{a, d\}$ is missing.

EXAMPLE 11.5

Let $G = (V, E)$ denote the graph in Fig. 11.16(a). The subgraphs in parts (b) and (c) of the figure are induced subgraphs of G. For the connected subgraph in part (b), $G_1 = \langle U_1 \rangle$ for $U_1 = \{b, c, d, e\}$. In like manner, the disconnected subgraph in part (c) is $G_2 = \langle U_2 \rangle$ for $U_2 = \{a, b, e, f\}$. Finally, G_3 in part (d) of Fig. 11.16 is a subgraph of G. But it is not an induced subgraph; the vertices c, e are in G_3, but the edge $\{c, e\}$ (of G) is not present.

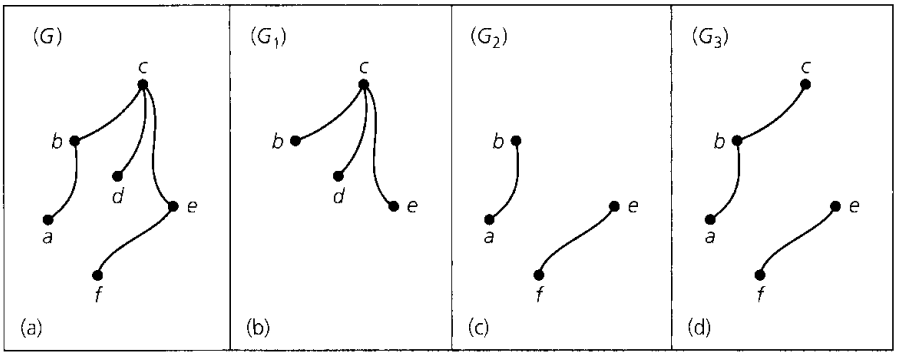


Figure 11.16

Another special type of subgraph comes about when a certain vertex or edge is deleted from the given graph. We formalize these ideas in the following definition.

Definition 11.10

Let v be a vertex in a directed or an undirected graph $G = (V, E)$. The subgraph of G denoted by $G - v$ has the vertex set $V_1 = V - \{v\}$ and the edge set $E_1 \subseteq E$, where E_1 contains all the edges in E except for those that are incident with the vertex v. (Hence $G - v$ is the subgraph of G induced by V_1.)

In a similar way, if e is an edge of a directed or an undirected graph $G = (V, E)$, we obtain the subgraph $G - e = (V_1, E_1)$ of G, where the set of edges $E_1 = E - \{e\}$, and the vertex set is unchanged (that is, $V_1 = V$).

EXAMPLE 11.6

Let $G = (V, E)$ be the undirected graph in Fig. 11.17(a). Part (b) of this figure is the subgraph G_1 (of G), where $G_1 = G - c$. It is also the subgraph of G induced by the set of vertices $U_1 = \{a, b, d, f, g, h\}$, so $G_1 = \langle V - \{c\} \rangle = \langle U_1 \rangle$. In part (c) of Fig. 11.17 we find the subgraph G_2 of G, where $G_2 = G - e$ for e the edge $\{c, d\}$. The result in Fig. 11.17(d) shows how the ideas in Definition 11.10 can be extended to the deletion of more than one vertex (edge). We may represent this subgraph of G as $G_3 = (G - b) - f = (G - f) - b = G - \{b, f\} = \langle U_3 \rangle$, for $U_3 = \{a, c, d, g, h\}$.

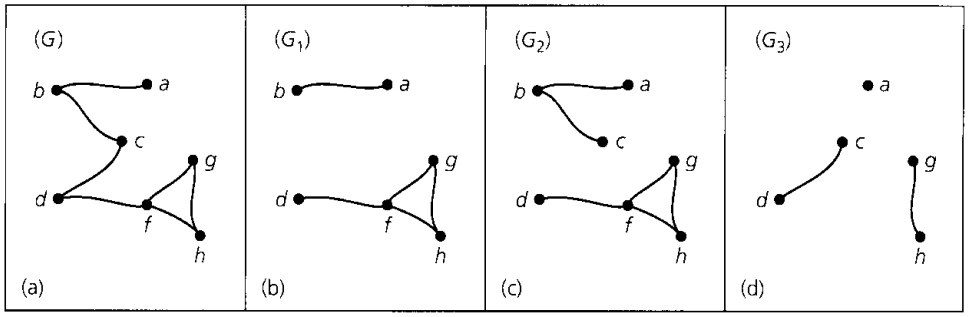


Figure 11.17

The idea of a subgraph gives us a way to develop the complement of an undirected loop-free graph. Before doing so, however, we define a type of graph that is maximal in size for a given number of vertices.

Definition 11.11 Let V be a set of n vertices. The *complete graph* on V, denoted K_n, is a loop-free undirected graph, where for all $a, b \in V$, $a \neq b$, there is an edge $\{a, b\}$.

Figure 11.18 provides the complete graphs K_n, for $1 \leq n \leq 4$. We shall realize, when we examine the idea of graph isomorphism, that these are the only possible complete graphs for the given number of vertices.

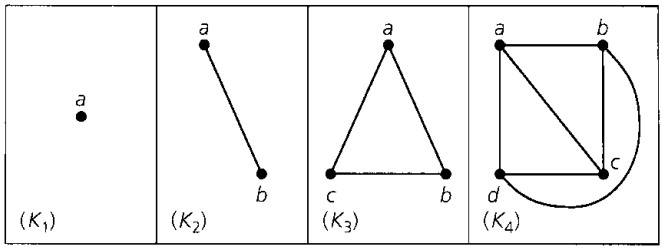


Figure 11.18

In determining the complement of a set in Chapter 3, we needed to know the universal set under consideration. The complete graph plays a role similar to a universal set.

Definition 11.12 Let G be a loop-free undirected graph on n vertices. The *complement of* G, denoted $\overline{G}$, is the subgraph of K_n consisting of the n vertices in G and all edges that are not in G. (If $G = K_n$, $\overline{G}$ is a graph consisting of n vertices and no edges. Such a graph is called a *null* graph.)

Figure 11.19(a) shows an undirected graph on four vertices. Its complement is shown in part (b) of the figure. In the complement, vertex a is isolated.

Once again we have reached a point where many new ideas have been defined. To demonstrate why some of these ideas are important, we apply them now to the solution of an interesting puzzle.

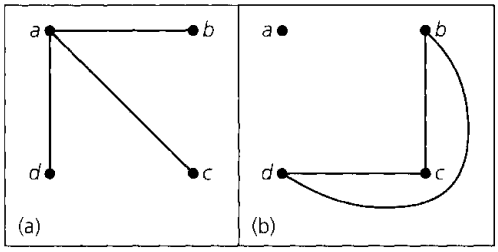


Figure 11.19

EXAMPLE 11.7

Instant Insanity. The game of Instant Insanity is played with four cubes. Each of the six faces on a cube is painted with one of the colors red (R), white (W), blue (B), or yellow (Y). The object of the game is to place the cubes in a column of four such that all four (different) colors appear on each of the four sides of the column.

Consider the cubes in Fig. 11.20 and number them as shown. (These cubes are only one example of this game. Many others exist.) First we shall estimate the number of arrangements that are possible here. If we wish to place cube 1 at the bottom of the column, there are at most three different ways in which we can do this. In Fig. 11.20 cube 1 is unfolded, and we see that it makes no difference whether we place the red face on the table or the opposite white face on the table. We are concerned only with the other four faces at the base of our column. With three pairs of opposite faces there will be at most three ways to place the *first* cube for the base of the column. Now consider cube 2. Although some colors are repeated, no pair of opposite faces has the same color. Hence we have six ways to place the second cube on top of the first. We can then rotate the second cube without changing either the face on the top of the first cube or the face on the bottom of the second cube. With four possible rotations we may place the second cube on top of the first in as many as 24 different ways. Continuing the argument, we find that there can be as many as $(3)(24)(24)(24) = 41{,}472$ possibilities to consider. And there may not even be a solution!

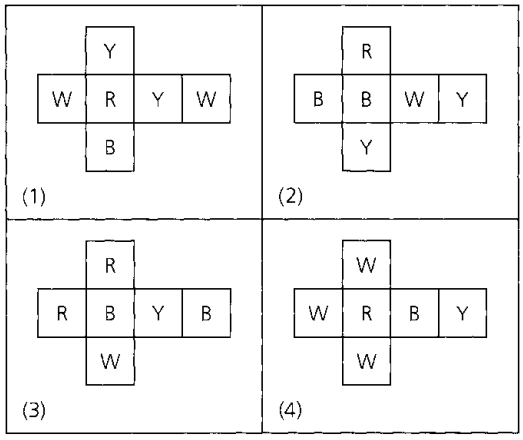


Figure 11.20

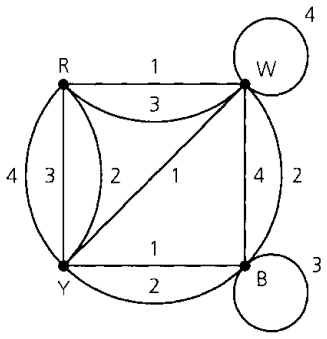


Figure 11.21

In solving this puzzle we realize that it is difficult to keep track of (1) colors on opposite faces of cubes and (2) columns of colors. A graph (actually a labeled multigraph) helps us to visualize the situation. In Fig. 11.21 we have a graph on four vertices R, W, B, and Y. As we consider each cube, we examine its three pairs of opposite faces. For example, cube

1 has a pair of opposite faces painted yellow and blue, so we draw an edge connecting Y and B and label it 1 (for cube 1). The other two edges in the figure that are labeled with 1 account for the pairs of opposite faces that are white and yellow, and red and white. Doing likewise for the other cubes, we arrive at the graph in the figure. A loop, such as the one at B, with label 3, indicates a pair of opposite faces with the same color (for cube 3).

In the graph we see a total of 12 edges falling into four sets of 3, according to the labels for the cubes. At each vertex the number of edges incident to (or from) the vertex counts the number of faces on the four cubes that have that color. (We count a loop twice.) Hence Fig. 11.21 tells us that for our four cubes we have five red faces, seven white ones, six blue ones, and six that are yellow.

With the four cubes stacked in a column, we examine two opposite sides of the column. This arrangement gives us four edges in the graph of Fig. 11.21, where each label appears once. Since each color is to appear only once on a side of the column, each color must appear twice as an endpoint of these four edges. If we can accomplish the same result for the other two sides of the column, we have solved the puzzle. In Fig. 11.22(a) we see that each side in one pair of opposite sides of our column has the four colors if the cubes are arranged according to the information provided by the subgraph shown there. However, to accomplish this for the other two sides of the column also, we need a second such subgraph that doesn't use any edge in part (a). In this case a second such subgraph does exist, as shown in part (b) of the figure.

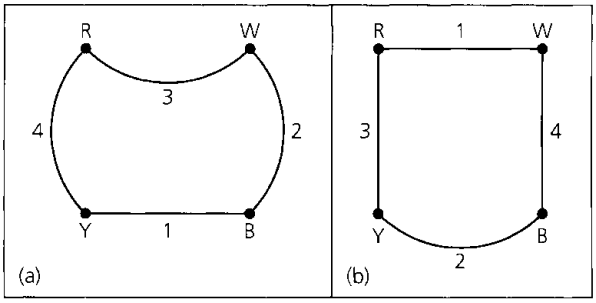


Figure 11.22

Figure 11.23 shows how to arrange the cubes as indicated by the subgraphs in Fig. 11.22.

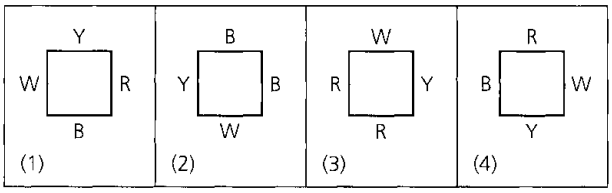


Figure 11.23

In general, for any four cubes we construct a labeled multigraph and try to find two subgraphs where (1) each subgraph contains all four vertices, and four edges, one for each label; (2) in each subgraph, each vertex is incident with exactly two edges (a loop is counted twice); and (3) no (labeled) edge of the labeled multigraph appears in both subgraphs.

Now we turn to the second question posed at the start of the section.

Parts (a) and (b) of Fig. 11.24 show two undirected graphs on four vertices. Since straight edges and curved edges are considered the same here, each graph represents six adjacent pairs of vertices. In fact, we probably feel that these graphs are both examples of the graph K_4. We make this feeling mathematically rigorous in the following definition.

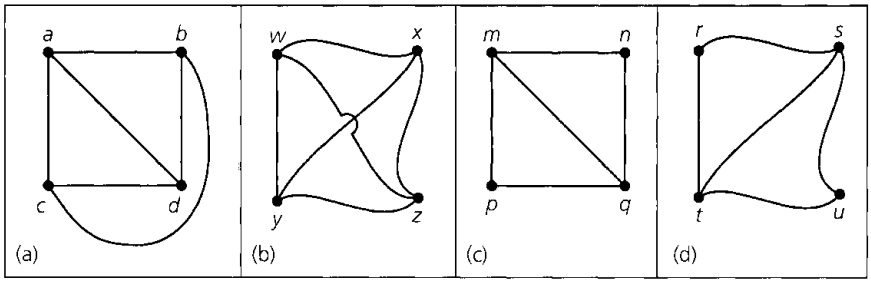


Figure 11.24

Definition 11.13 Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two undirected graphs. A function $f: V_1 \to V_2$ is called a *graph isomorphism* if (a) f is one-to-one and onto, and (b) for all $a, b \in V_1$, $\{a, b\} \in E_1$ if and only if $\{f(a), f(b)\} \in E_2$. When such a function exists, G_1 and G_2 are called *isomorphic graphs*.

The vertex correspondence of a graph isomorphism preserves adjacencies. Since which pairs of vertices are adjacent and which are not is the only essential property of an undirected graph, in this way the structure of the graphs is preserved.

For the graphs in parts (a) and (b) of Fig. 11.24 the function f defined by

$$f(a) = w, \qquad f(b) = x, \qquad f(c) = y, \qquad f(d) = z$$

provides an isomorphism. [In fact, any one-to-one correspondence between $\{a, b, c, d\}$ and $\{w, x, y, z\}$ will be an isomorphism because both of the given graphs are complete graphs. This would also be true if each of the given graphs had only four isolated vertices (and no edges).] Consequently, as far as (graph) structure is concerned, these graphs are considered the same — each is (isomorphic to) the complete graph K_4.

For the graphs in parts (c) and (d) of Fig. 11.24 we need to be a little more careful. The function g defined by

$$g(m) = r, \qquad g(n) = s, \qquad g(p) = t, \qquad g(q) = u$$

is one-to-one and onto (for the given vertex sets). However, although $\{m, q\}$ is an edge in the graph of part (c), $\{g(m), g(q)\} = \{r, u\}$ is not an edge in the graph of part (d). Consequently, the function g does *not* define a graph isomorphism. To maintain the correspondence of edges, we consider the one-to-one onto function h where

$$h(m) = s, \qquad h(n) = r, \qquad h(p) = u, \qquad h(q) = t.$$

In this case we have the edge correspondences

$$\{m, n\} \leftrightarrow \{h(m), h(n)\} = \{s, r\}, \qquad \{n, q\} \leftrightarrow \{h(n), h(q)\} = \{r, t\},$$
$$\{m, p\} \leftrightarrow \{h(m), h(p)\} = \{s, u\}, \qquad \{p, q\} \leftrightarrow \{h(p), h(q)\} = \{u, t\},$$
$$\{m, q\} \leftrightarrow \{h(m), h(q)\} = \{s, t\},$$

so h is a graph isomorphism. [We also notice how, for example, the cycle $m \to n \to q \to m$ corresponds with the cycle $s\ (= h(m)) \to r\ (= h(n)) \to t\ (= h(q)) \to s\ (= h(m))$.]

Finally, since the graph in part (a) of Fig. 11.24 has six edges and that in part (c) has only five edges, these two graphs cannot be isomorphic.

Now let us examine the idea of graph isomorphism in a more difficult situation.

EXAMPLE 11.8

In Fig. 11.25 we have two graphs, each on ten vertices. Unlike the graphs in Fig. 11.24, it is not immediately apparent whether or not these graphs are isomorphic.

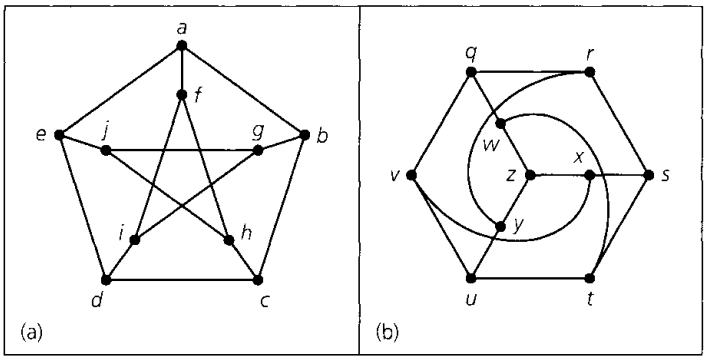


Figure 11.25

One finds that the correspondence given by

$$a \to q \quad c \to u \quad e \to r \quad g \to x \quad i \to z$$
$$b \to v \quad d \to y \quad f \to w \quad h \to t \quad j \to s$$

preserves all adjacencies. For example, $\{f, h\}$ is an edge in graph (a) with $\{w, t\}$ the corresponding edge in graph (b). But how did we come up with the correspondence? The following discussion provides some clues.

We note that because an isomorphism preserves adjacencies, it preserves graph substructures such as paths and cycles. In graph (a) the edges $\{a, f\}$, $\{f, i\}$, $\{i, d\}$, $\{d, e\}$, and $\{e, a\}$ constitute a cycle of length 5. Hence we must preserve this as we try to find an isomorphism. One possibility for the corresponding edges in graph (b) is $\{q, w\}$, $\{w, z\}$, $\{z, y\}$, $\{y, r\}$, and $\{r, q\}$, which also provides a cycle of length 5. (A second possible choice is given by the edges in the cycle $y \to r \to s \to t \to u \to y$.) In addition, starting at vertex a in graph (a), we find a path that will "visit" each vertex only once. We express this path by $a \to f \to h \to c \to b \to g \to j \to e \to d \to i$. For the graphs to be isomorphic there must be a corresponding path in graph (b). Here the path described by $q \to w \to t \to u \to v \to x \to s \to r \to y \to z$ is the counterpart.

These are some of the ideas we can use to try to develop an isomorphism and determine whether two graphs are isomorphic. Other considerations will be discussed throughout the chapter. However, there is no simple, foolproof method — especially when we are confronted with larger graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, where $|V_1| = |V_2|$ and $|E_1| = |E_2|$.

We close this section with one more example involving graph isomorphism.

EXAMPLE 11.9

Each of the two graphs in Fig. 11.26 has six vertices and nine edges. Therefore it is reasonable to ask whether they are isomorphic.

In graph (a), vertex a is adjacent to two other vertices of the graph. Consequently, if we try to construct an isomorphism between these graphs, we should associate vertex a with a comparable vertex in graph (b), say vertex u. A similar situation exists for vertex d and either vertex x or vertex z. But no matter which of the vertices x or z we use, there remains one vertex in graph (b) that is adjacent to two other vertices. And there is no other such vertex in graph (a) to continue our one-to-one structure preserving correspondence. Consequently, these graphs are not isomorphic.

Furthermore, in graph (b) it is possible to start at any vertex and find a circuit that includes every edge of the graph. For example, if we start at vertex u, the circuit $u \to w \to v \to y \to w \to z \to y \to x \to v \to u$ exhibits this property. This does not happen in graph (a) where the only trails that include each edge start at either b or f and then terminate at f or b, respectively.

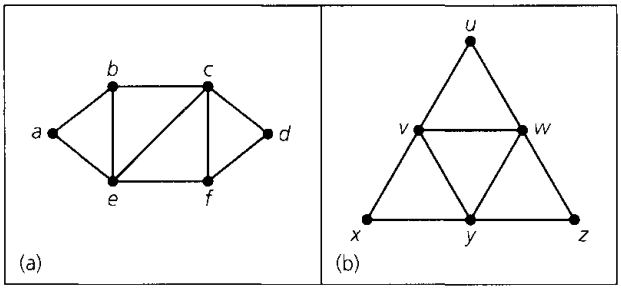


Figure 11.26

EXERCISES 11.2

1. Let G be the undirected graph in Fig. 11.27(a).

   a) How many connected subgraphs of G have four vertices and include a cycle?

   b) Describe the subgraph G_1 (of G) in part (b) of the figure first, as an induced subgraph and second, in terms of deleting a vertex of G.

   c) Describe the subgraph G_2 (of G) in part (c) of the figure first, as an induced subgraph and second, in terms of the deletion of vertices of G.

   d) Draw the subgraph of G induced by the set of vertices $U = \{b, c, d, f, i, j\}$.

   e) For the graph G, let the edge $e = \{c, f\}$. Draw the subgraph $G - e$.

2. **a)** Let $G = (V, E)$ be an undirected graph, with $G_1 = (V_1, E_1)$ a subgraph of G. Under what condition(s) is G_1 *not* an induced subgraph of G?

   b) For the graph G in Fig. 11.27(a), find a subgraph that is not an induced subgraph.

3. **a)** How many spanning subgraphs are there for the graph G in Fig. 11.27(a)?

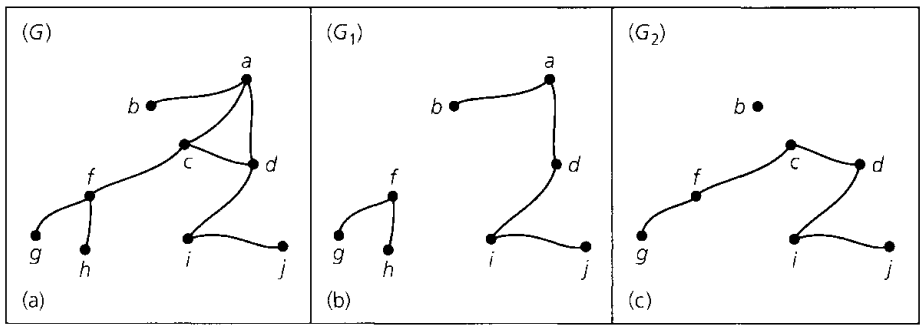


Figure 11.27

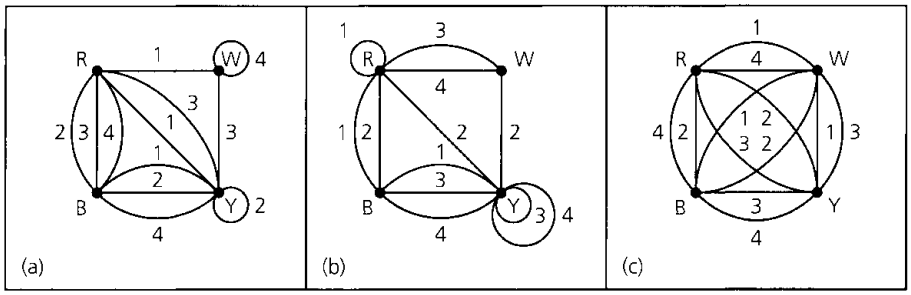


Figure 11.28

b) How many connected spanning subgraphs are there in part (a)?

c) How many of the spanning subgraphs in part (a) have vertex a as an isolated vertex?

4. If $G = (V, E)$ is an undirected graph, how many spanning subgraphs of G are also induced subgraphs?

5. Let $G = (V, E)$ be an undirected graph, where $|V| \geq 2$. If every induced subgraph of G is connected, can we identify the graph G?

6. Find all (loop-free) nonisomorphic undirected graphs with four vertices. How many of these graphs are connected?

7. Each of the labeled multigraphs in Fig. 11.28 arises in the analysis of a set of four blocks for the game of Instant Insanity. In each case determine a solution to the puzzle, if possible.

8. a) How many paths of length 4 are there in the complete graph K_7? (Remember that a path such as $v_1 \to v_2 \to v_3 \to v_4 \to v_5$ is considered to be the same as the path $v_5 \to v_4 \to v_3 \to v_2 \to v_1$.)

b) Let $m, n \in \mathbf{Z}^+$ with $m < n$. How many paths of length m are there in the complete graph K_n?

9. For each pair of graphs in Fig. 11.29, determine whether or not the graphs are isomorphic.

10. Let G be an undirected (loop-free) graph with v vertices and e edges. How many edges are there in $\overline{G}$?

11. a) If G_1, G_2 are (loop-free) undirected graphs, prove that G_1, G_2 are isomorphic if and only if $\overline{G}_1$, $\overline{G}_2$ are isomorphic.

b) Determine whether the graphs in Fig. 11.30 are isomorphic.

12. a) Let G be an undirected graph with n vertices. If G is isomorphic to its own complement $\overline{G}$, how many edges must G have? (Such a graph is called *self-complementary*.)

b) Find an example of a self-complementary graph on four vertices and one on five vertices.

c) If G is a self-complementary graph on n vertices, where $n > 1$, prove that $n = 4k$ or $n = 4k + 1$, for some $k \in \mathbf{Z}^+$.

13. Let G be a cycle on n vertices. Prove that G is self-complementary if and only if $n = 5$.

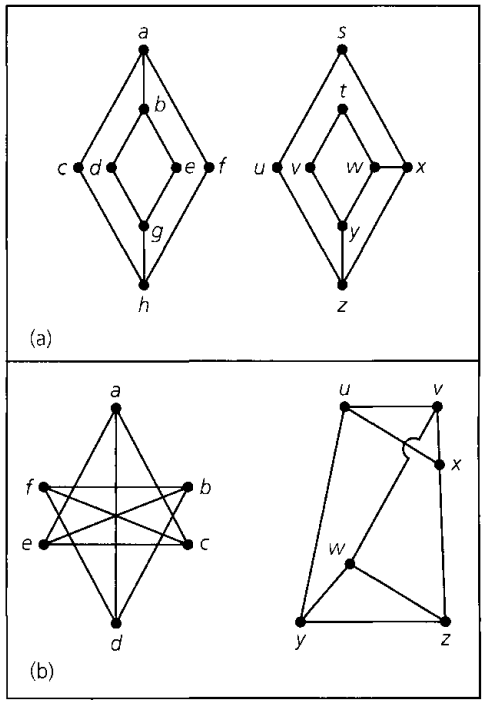


Figure 11.29

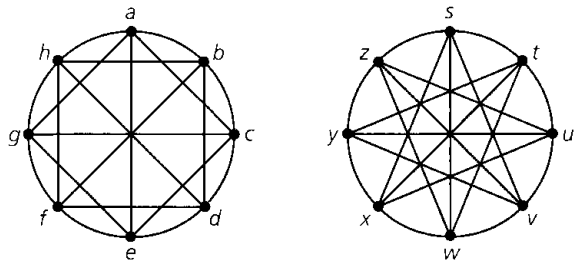


Figure 11.30

14. a) Find a graph G where both G and $\overline{G}$ are connected.

b) If G is a graph on n vertices, for $n \geq 2$, and G is not connected, prove that $\overline{G}$ is connected.

15. **a)** Extend Definition 11.13 to directed graphs.

b) Determine whether the directed graphs in Fig. 11.31 are isomorphic.

16. **a)** How many subgraphs $H = (V, E)$ of K_6 satisfy $|V| = 3$? (If two subgraphs are isomorphic but have different vertex sets, consider them distinct.)

b) How many subgraphs $H = (V, E)$ of K_6 satisfy $|V| = 4$?

c) How many subgraphs does K_6 have?

d) For $n \geq 3$, how many subgraphs does K_n have?

17. Let v, w be two vertices in K_n, $n \geq 3$. How many walks of length 3 are there from v to w?

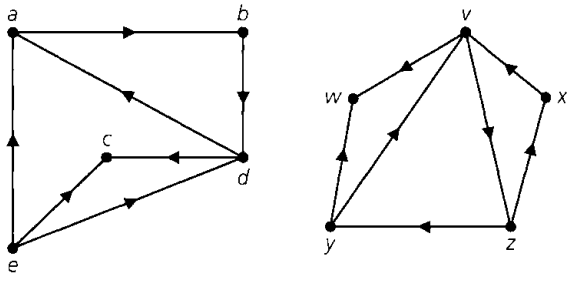


Figure 11.31

11.3 Vertex Degree: Euler Trails and Circuits

In Example 11.9 the number of edges incident with a vertex was used to show that two undirected graphs were not isomorphic. We now find this idea even more helpful.

Definition 11.14 Let G be an undirected graph or multigraph. For each vertex v of G, the *degree of* v, written $\deg(v)$, is the number of edges in G that are incident with v. Here a loop at a vertex v is considered as two incident edges for v.

EXAMPLE 11.10 For the graph in Fig. 11.32, $\deg(b) = \deg(d) = \deg(f) = \deg(g) = 2$, $\deg(c) = 4$, $\deg(e) = 0$, and $\deg(h) = 1$. For vertex a we have $\deg(a) = 3$ because we count a loop twice. Since h has degree 1, it is called a *pendant* vertex.

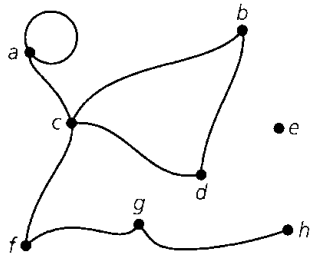


Figure 11.32

Using the idea of vertex degree, we have the following result.

THEOREM 11.2 If $G = (V, E)$ is an undirected graph or multigraph, then $\sum_{v \in V} \deg(v) = 2|E|$.

Proof: As we consider each edge $\{a, b\}$ in graph G, we find that the edge contributes a count of 1 to each of $\deg(a)$, $\deg(b)$, and consequently a count of 2 to $\sum_{v \in V} \deg(v)$. Thus $2|E|$ accounts for $\deg(v)$, for all $v \in V$, and $\sum_{v \in V} \deg(v) = 2|E|$.

This theorem provides some insight into the number of odd-degree vertices that can exist in a graph.

COROLLARY 11.1 For any undirected graph or multigraph, the number of vertices of odd degree must be even.

Proof: We leave the proof for the reader.

We apply Theorem 11.2 in the following example.

EXAMPLE 11.11

An undirected graph (or multigraph) where each vertex has the same degree is called a *regular* graph. If $\deg(v) = k$ for all vertices v, then the graph is called *k-regular*. Is it possible to have a 4-regular graph with 10 edges?

From Theorem 11.2, $2|E| = 20 = 4|V|$, so we have five vertices of degree 4. Figure 11.33 provides two nonisomorphic examples that satisfy the requirements.

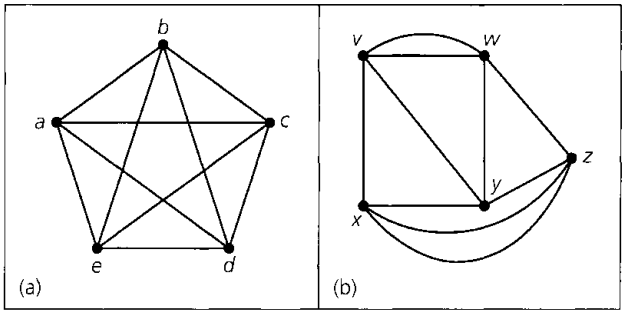


Figure 11.33

If we want each vertex to have degree 4, with 15 edges in the graph, we find that $2|E| = 30 = 4|V|$, from which it follows that no such graph is possible.

Our next example introduces a regular graph that arises in the study of computer architecture.

EXAMPLE 11.12

The Hypercube. In order to build a parallel computer one needs to have multiple CPUs (central processing units), where each such processor works on part of a problem. But often we cannot actually decompose a problem completely, so at some point the processors (each with its own memory) have to be able to communicate with one another.

We envisage this situation as follows. The accumulated data for a given problem are taken from a central storage location and divided up among the processors. The processors go through a phase where each computes on its own for a certain period of time and then some intercommunication takes place. Then the processors return to computing on their own and continue back and forth between operating individually and communicating with one another. This situation adequately describes how parallel algorithms work in practice.

To model the communication between the processors we use a loop-free connected undirected graph where each processor is assigned a vertex. When two processors, say p_1, p_2, are able to communicate directly with one another we draw the edge $\{p_1, p_2\}$ to represent this (line of) possible communication. How can we decide on a model (that is, a graph) to speed up the processing time? The complete graph (on all of our processors as vertices)

would be ideal — but prohibitively expensive because of all the necessary connections. On the other hand, one can connect n processors along a path with $n - 1$ edges or on a cycle with n edges. Another possible model is a *grid* (or, *mesh*) graph, examples of which are shown in Fig. 11.34.

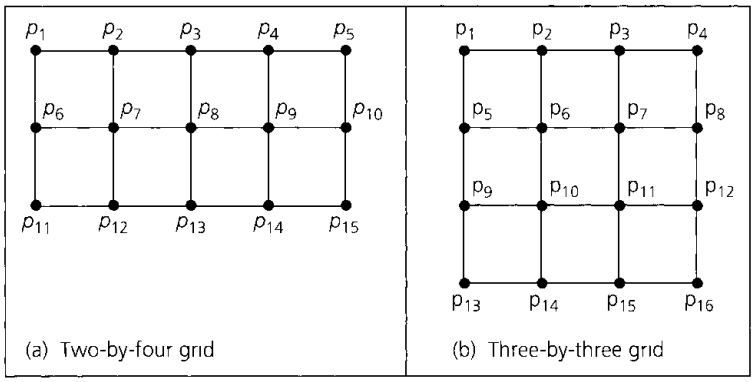


Figure 11.34

But in these last three models the distances (as measured by the number of edges in the shortest paths) between pairs of processors get longer and longer as the number of processors increases. A compromise that weighs the number of edges (direct connections) against the distance between pairs of vertices (processors) is embodied in the regular graph called the *hypercube*.

For $n \in \mathbf{N}$, the n-dimensional hypercube (or *n-cube*) is denoted by Q_n. It is a loop-free connected undirected graph with 2^n vertices. For $n \geq 1$, these vertices are labeled by the 2^n n-bit sequences representing 0, 1, 2, . . . , $2^n - 1$. For instance, Q_3 has eight vertices — labeled 000, 001, 010, 011, 100, 101, 110, and 111. Two vertices v_1, v_2 of Q_n are joined by the edge $\{v_1, v_2\}$ when the binary labels for v_1, v_2 differ in exactly one position. Then for any vertices u, w in Q_n there is a shortest path of length d, when d is the number of positions where the binary labels for u, w differ. [This insures that Q_n is connected.]

Figure 11.35 shows Q_n for $n = 0, 1, 2, 3$. In general, for $n \geq 0$, Q_{n+1} can be constructed recursively from two copies of Q_n as follows. Prefix the vertex labels of one copy of Q_n with 0 (call the result $Q_{0,n}$) and those of the other copy with 1 (call this result $Q_{1,n}$). For x in $Q_{0,n}$ and y in $Q_{1,n}$ draw the edge $\{x, y\}$ if the (newly prefixed) binary labels for x, y differ only in the first (newly prefixed) position. The case for $n = 3$ (so $n + 1 = 4$) is demonstrated in Fig. 11.36. The blue edges are the new edges described above for constructing Q_4 from two copies of Q_3.

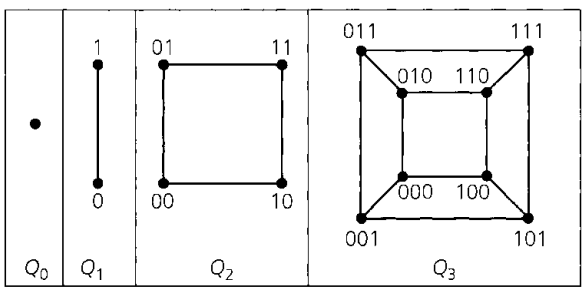


Figure 11.35

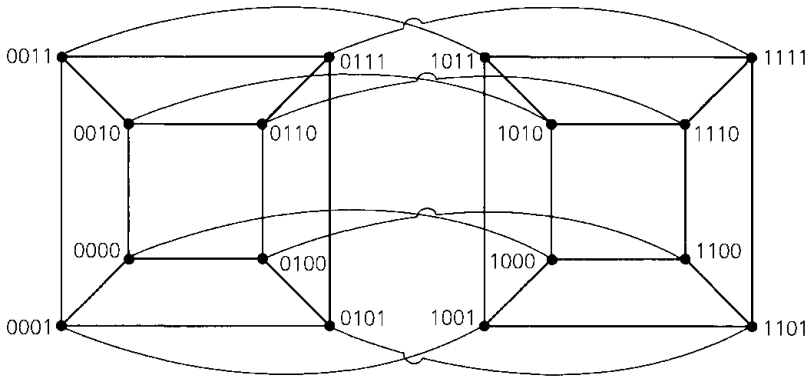


Figure 11.36

In summary, we reiterate that for $n \in \mathbf{N}$, the hypercube Q_n is an n-regular loop-free undirected graph with 2^n vertices. Further, it is connected with the distance between any two vertices at most n. From Theorem 11.2 it follows that Q_n has $(1/2)n2^n = n2^{n-1}$ edges. [Referring back to Example 10.33, we find that $n2^{n-1}$ is likewise the number of edges for the Hasse diagram of the partial order $(\mathcal{P}(X_n), \subseteq)$, where $X_n = \{1, 2, 3, \ldots, n\}$ and $\mathcal{P}(X_n)$ is the power set of X_n. This is no mere coincidence! If we use the Gray code of Example 3.9 to label the vertices of this Hasse diagram, we find we have the hypercube Q_n.]

Finally, note that in Q_4 there are 16 vertices (processors) and the longest distance between vertices is 4. Contrast this with the grids in Fig. 11.34, where there are 15 vertices in part (a) and 16 in part (b)—yet the longest distance is 6 in both grids.

We turn now to the reason why Euler developed the idea of the degree of a vertex: to solve the problem dealing with the seven bridges of Königsberg.

EXAMPLE 11.13

The Seven Bridges of Königsberg. During the eighteenth century, the city of Königsberg (in East Prussia) was divided into four sections (including the island of Kneiphof) by the Pregel River. Seven bridges connected these regions, as shown in Fig. 11.37(a). It was said that residents spent their Sunday walks trying to find a way to walk about the city so as to cross each bridge exactly once and then return to the starting point.

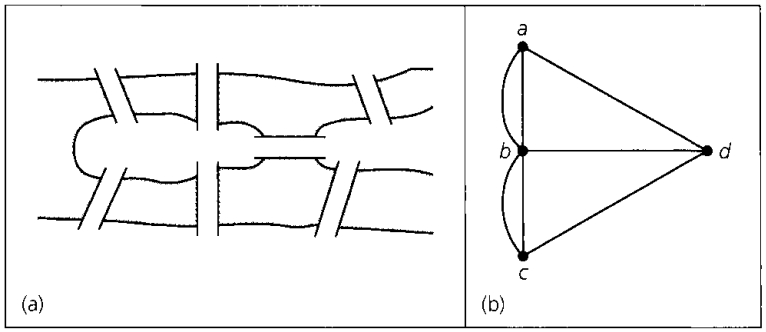


Figure 11.37

In order to determine whether or not such a circuit existed, Euler represented the four sections of the city and the seven bridges by the multigraph shown in Fig. 11.37(b). Here

he found four vertices with $\deg(a) = \deg(c) = \deg(d) = 3$ and $\deg(b) = 5$. He also found that the existence of such a circuit depended on the number of vertices of odd degree in the graph.

Before proving the general result, we give the following definition.

Definition 11.15 Let $G = (V, E)$ be an undirected graph or multigraph with no isolated vertices. Then G is said to have an *Euler circuit* if there is a circuit in G that traverses every edge of the graph exactly once. If there is an open trail from a to b in G and this trail traverses each edge in G exactly once, the trail is called an *Euler trail.*

The problem of the seven bridges is now settled as we characterize the graphs that have an Euler circuit.

THEOREM 11.3 Let $G = (V, E)$ be an undirected graph or multigraph with no isolated vertices. Then G has an Euler circuit if and only if G is connected and every vertex in G has even degree.

Proof: If G has an Euler circuit, then for all $a, b \in V$ there is a trail from a to b — namely, that part of the circuit that starts at a and terminates at b. Therefore, it follows from Theorem 11.1 that G is connected.

Let s be the starting vertex of the Euler circuit. For any other vertex v of G, each time the circuit comes to v it then departs from the vertex. Thus the circuit has traversed either two (new) edges that are incident with v or a (new) loop at v. In either case a count of 2 is contributed to $\deg(v)$. Since v is not the starting point and each edge incident to v is traversed only once, a count of 2 is obtained each time the circuit passes through v, so $\deg(v)$ is even. As for the starting vertex s, the first edge of the circuit must be distinct from the last edge, and because any other visit to s results in a count of 2 for $\deg(s)$, we have $\deg(s)$ even.

Conversely, let G be connected with every vertex of even degree. If the number of edges in G is 1 or 2, then G must be as shown in Fig. 11.38. Euler circuits are immediate in these cases. We proceed now by induction and assume the result true for all situations where there are fewer than n edges. If G has n edges, select a vertex s in G as a starting point to build an Euler circuit. The graph (or multigraph) G is connected and each vertex has even degree, so we can at least construct a circuit C containing s. (Verify this by considering the longest trail in G that starts at s.) Should the circuit contain every edge of G, we are finished. If not, remove the edges of the circuit from G, making sure to remove any vertex that would become isolated. The remaining subgraph K has all vertices of even degree, but it may not be connected. However, each component of K is connected and will have an Euler circuit. (Why?) In addition, each of these Euler circuits has a vertex that is on C. Consequently, starting at s we travel on C until we arrive at a vertex s_1 that is on the Euler circuit of a

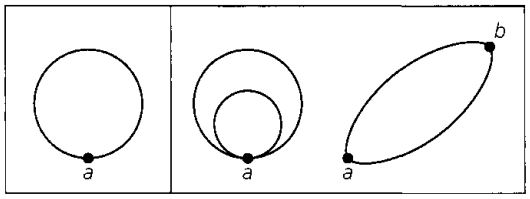


Figure 11.38

component C_1 of K. Then we traverse this Euler circuit and, returning to s_1, continue on C until we reach a vertex s_2 that is on the Euler circuit of component C_2 of K. Since the graph G is finite, as we continue this process we construct an Euler circuit for G.

Should G be connected and not have too many vertices of odd degree, we can at least find an Euler trail in G.

COROLLARY 11.2

If G is an undirected graph or multigraph with no isolated vertices, then we can construct an Euler trail in G if and only if G is connected and has exactly two vertices of odd degree.

Proof: If G is connected and a and b are the vertices of G that have odd degree, add an additional edge $\{a, b\}$ to G. We now have a graph G_1 that is connected and has every vertex of even degree. Hence G_1 has an Euler circuit C, and when the edge $\{a, b\}$ is removed from C, we obtain an Euler trail for G. (Thus the Euler trail starts at one of the vertices of odd degree and terminates at the other odd vertex.) We leave the details of the converse for the reader.

Returning now to the seven bridges of Königsberg, we realize that Fig. 11.37(b) is a connected multigraph, but it has four vertices of odd degree. Consequently, it has no Euler trail or Euler circuit.

Now that we have seen how the solution of an eighteenth-century problem led to the start of graph theory, is there a somewhat more contemporary context in which we might be able to apply what we have learned?

To answer this question (in the affirmative), we shall state the directed version of Theorem 11.3. But first we need to refine the concept of the degree of a vertex.

Definition 11.16

Let $G = (V, E)$ be a directed graph or multigraph. For each $v \in V$,

a) The *incoming*, or *in*, *degree* of v is the number of edges in G that are incident into v, and this is denoted by $id(v)$.

b) The *outgoing*, or *out*, *degree* of v is the number of edges in G that are incident from v, and this is denoted by $od(v)$.

For the case where the directed graph or multigraph contains one or more loops, each loop at a given vertex v contributes a count of 1 to each of $id(v)$ and $od(v)$.

The concepts of the in degree and the out degree for vertices now lead us to the following theorem.

THEOREM 11.4

Let $G = (V, E)$ be a directed graph or multigraph with no isolated vertices. The graph G has a directed Euler circuit if and only if G is connected and $id(v) = od(v)$ for all $v \in V$.

Proof: The proof of this theorem is left for the reader.

At this time we consider an application of Theorem 11.4. This example is based on a telecommunication problem given by C. L. Liu on pages 176–178 of reference [23].

EXAMPLE 11.14

In Fig. 11.39(a) we have the surface of a rotating drum that is divided into eight sectors of equal area. In part (b) of the figure we have placed conducting (shaded sectors and inner circle) and nonconducting (unshaded sectors) material on the drum. When the three terminals (shown in the figure) make contact with the three designated sectors, the nonconducting material results in no flow of current and a 1 appears on the display of a digital device. For the sectors with the conducting material, a flow of current takes place and a 0 appears on the display in each case. If the drum were rotated 45 degrees (clockwise), the screen would read 110 (from top to bottom). So we can obtain at least two (namely, 100 and 110) of the eight binary representations from 000 (for 0) to 111 (for 7). But can we represent all eight of them as the drum continues to rotate? And could we extend the problem to the 16 four-bit binary representations from 0000 through 1111, and perhaps generalize the results even further?

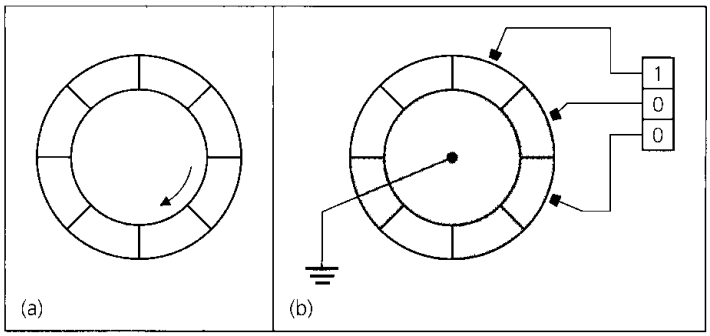


Figure 11.39

To answer the question for the problem in the figure, we construct a directed graph $G = (V, E)$, where $V = \{00, 01, 10, 11\}$ and E is constructed as follows: If b_1b_2, $b_2b_3 \in V$, draw the edge (b_1b_2, b_2b_3). This results in the directed graph of Fig. 11.40(a), where $|E| = 8$. We see that this graph is connected and that for all $v \in V$, $id(v) = od(v)$. Consequently, by Theorem 11.4, it has a directed Euler circuit. One such circuit is given by

$$\underset{\curvearrowright}{10} \xrightarrow{100} 00 \xrightarrow{000} 00 \xrightarrow{001} 01 \xrightarrow{010} 10 \xrightarrow{101} 01 \xrightarrow{011} 11 \xrightarrow{111} \underset{\curvearrowright}{11}$$

$$110$$

Here the label on each edge $e = (a, c)$, as shown in part (b) of Fig. 11.40, is the three-bit sequence $x_1x_2x_3$, where $a = x_1x_2$ and $c = x_2x_3$. Since the vertices of G are the four distinct two-bit sequences 00, 01, 10, and 11, the labels on the eight edges of G determine the eight distinct three-bit sequences. Also, any two consecutive edge labels in the Euler circuit are of the form $y_1y_2y_3$ and $y_2y_3y_4$.

Starting with the edge label 100, in order to get the next label, 000, we concatenate the last bit in 000, namely 0, to the string 100. The resulting string 1000 then provides 100 (1000) and 000 (1000). The next edge label is 001, so we concatenate the 1 (the last bit in 001) to our present string 1000 and get 10001, which provides the three distinct three-bit sequences 100 (10001), 000 (10001), and 001 (10001). Continuing in this way, we arrive at the eight-bit sequence 10001011 (where the last 1 is *wrapped around*), and these eight bits are then arranged in the sectors of the rotating drum as in Fig. 11.41. It is from this figure that the result in Fig. 11.39(b) is obtained. And as the drum in Fig. 11.39(b) rotates, all of the eight three-bit sequences 100, 110, 111, 011, 101, 010, 001, and 000 are obtained.

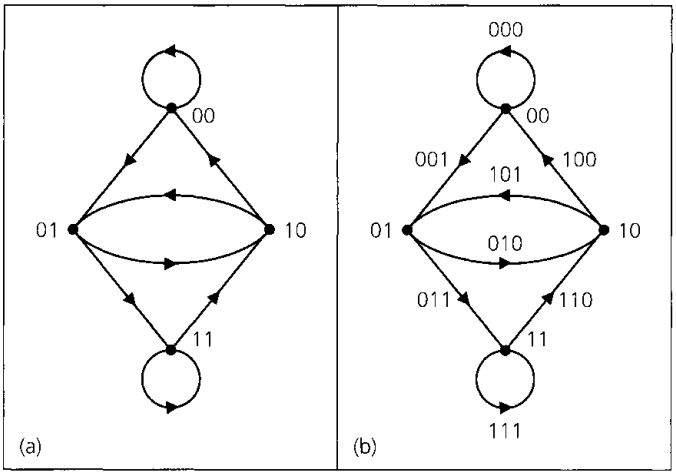


Figure 11.40

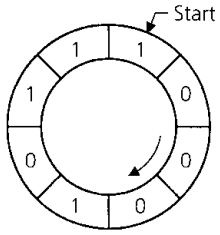


Figure 11.41

In closing this section, we wish to call the reader's attention to reference [24] by Anthony Ralston. This article is a good source for more ideas and generalizations related to the problem discussed in Example 11.14.

EXERCISES 11.3

1. Determine $|V|$ for the following graphs or multigraphs G.

a) G has nine edges and all vertices have degree 3.

b) G is regular with 15 edges.

c) G has 10 edges with two vertices of degree 4 and all others of degree 3.

2. If $G = (V, E)$ is a connected graph with $|E| = 17$ and $\deg(v) \geq 3$ for all $v \in V$, what is the maximum value for $|V|$?

3. Let $G = (V, E)$ be a connected undirected graph.

a) What is the largest possible value for $|V|$ if $|E| = 19$ and $\deg(v) \geq 4$ for all $v \in V$?

b) Draw a graph to demonstrate each possible case in part (a).

4. a) Let $G = (V, E)$ be a loop-free undirected graph, where $|V| = 6$ and $\deg(v) = 2$ for all $v \in V$, Up to isomorphism how many such graphs G are there?

b) Answer part (a) for $|V| = 7$.

c) Let $G_1 = (V_1, E_1)$ be a loop-free undirected 3-regular graph with $|V_1| = 6$. Up to isomorphism how many such graphs G_1 are there?

d) Answer part (c) for $|V_1| = 7$ and G_1 4-regular.

e) Generalize the results in parts (c) and (d).

5. Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be the loop-free undirected connected graphs in Fig. 11.42.

a) Determine $|V_1|$, $|E_1|$, $|V_2|$, and $|E_2|$.

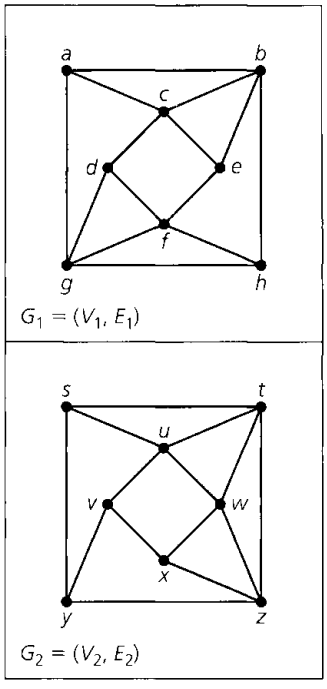


Figure 11.42

b) Find the degree of each vertex in V_1. Do likewise for each vertex in V_2.

c) Are the graphs G_1 and G_2 isomorphic?

6. Let $V = \{a, b, c, d, e, f\}$. Draw three nonisomorphic loop-free undirected graphs $G_1 = (V, E_1)$, $G_2 = (V, E_2)$, and

$G_3 = (V, E_3)$, where, in all three graphs, we have $\deg(a) = 3$, $\deg(b) = \deg(c) = 2$, and $\deg(d) = \deg(e) = \deg(f) = 1$.

7. **a)** How many different paths of length 2 are there in the undirected graph G in Fig. 11.43?

b) Let $G = (V, E)$ be a loop-free undirected graph, where $V = \{v_1, v_2, \ldots, v_n\}$ and $\deg(v_i) = d_i$, for all $1 \leq i \leq n$. How many different paths of length 2 are there in G?

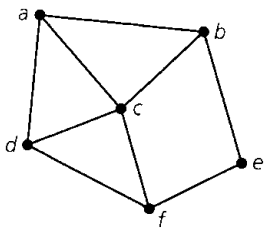


Figure 11.43

8. **a)** Find the number of edges in Q_8.

b) Find the maximum distance between pairs of vertices in Q_8. Give an example of one such pair that achieves this distance.

c) Find the length of a longest path in Q_8.

9. **a)** What is the dimension of the hypercube with 524,288 edges?

b) How many vertices are there for a hypercube with 4,980,736 edges?

10. For $n \in \mathbf{Z}^+$, how many distinct (though isomorphic) paths of length 2 are there in the n-dimensional hypercube Q_n?

11. Let $n \in \mathbf{Z}^+$, with $n \geq 9$. Prove that if the edges of K_n can be partitioned into subgraphs isomorphic to cycles of length 4 (where any two such cycles share no common edge), then $n = 8k + 1$ for some $k \in \mathbf{Z}^+$.

12. **a)** For $n \geq 2$, let V denote the vertices in Q_n. For $1 \leq k < \ell \leq n$, define the relation $\mathcal{R}$ on V as follows: If $w, x \in V$, then $w \mathcal{R} x$ if w and x have the same bit (0, or 1) in position k and the same bit (0, or 1) in position ℓ of their binary labels. [For example, if $n = 7$ and $k = 3$, $\ell = 6$, then 1100010 $\mathcal{R}$ 0000011.] Show that $\mathcal{R}$ is an equivalence relation. How many blocks are there for this equivalence relation? How many vertices are there in each block? Describe the subgraph of Q_n induced by the vertices in each block.

b) Generalize the results of part (a).

13. If G is an undirected graph with n vertices and e edges, let $\delta = \min_{v \in V}\{\deg(v)\}$ and let $\Delta = \max_{v \in V}\{\deg(v)\}$. Prove that $\delta \leq 2(e/n) \leq \Delta$.

14. Let $G = (V, E)$, $H = (V', E')$ be undirected graphs with $f: V \to V'$ establishing an isomorphism between the graphs. (a) Prove that $f^{-1}: V' \to V$ is also an isomorphism for G and H. (b) If $a \in V$, prove that $\deg(a)$ (in G) $= \deg(f(a))$ (in H).

15. For all $k \in \mathbf{Z}^+$ where $k \geq 2$, prove that there exists a loop-free connected undirected graph $G = (V, E)$, where $|V| = 2k$ and $\deg(v) = 3$ for all $v \in V$.

16. Prove that for each $n \in \mathbf{Z}^+$ there exists a loop-free connected undirected graph $G = (V, E)$, where $|V| = 2n$ and which has two vertices of degree i for every $1 \leq i \leq n$.

17. Complete the proofs of Corollaries 11.1 and 11.2.

18. Let k be a fixed positive integer and let $G = (V, E)$ be a loop-free undirected graph, where $\deg(v) \geq k$ for all $v \in V$. Prove that G contains a path of length k.

19. **a)** Explain why it is not possible to draw a loop-free connected undirected graph with eight vertices, where the degrees of the vertices are 1, 1, 1, 2, 3, 4, 5, and 7.

b) Give an example of a loop-free connected undirected multigraph with eight vertices, where the degrees of the vertices are 1, 1, 1, 2, 3, 4, 5, and 7.

20. **a)** Find an Euler circuit for the graph in Fig. 11.44.

b) If the edge $\{d, e\}$ is removed from this graph, find an Euler trail for the resulting subgraph.

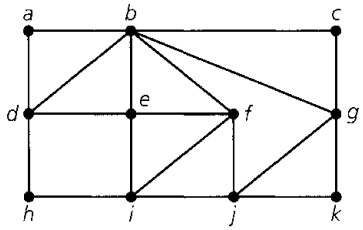


Figure 11.44

21. Determine the value(s) of n for which the complete graph K_n has an Euler circuit. For which n does K_n have an Euler trail but not an Euler circuit?

22. For the graph in Fig. 11.37(b), what is the smallest number of bridges that must be removed so that the resulting subgraph has an Euler trail but not an Euler circuit? Which bridge(s) should we remove?

23. When visiting a chamber of horrors, Paul and David try to figure out whether they can travel through the seven rooms and surrounding corridor of the attraction without passing through any door more than once. If they must start from the starred position in the corridor shown in Fig. 11.45, can they accomplish their goal?

24. Let $G = (V, E)$ be a directed graph, where $|V| = n$ and $|E| = e$. What are the values for $\sum_{v \in V} id(v)$ and $\sum_{v \in V} od(v)$?

25. **a)** Find the maximum length of a trail in

i) K_6 **ii)** K_8

iii) K_{10} **iv)** K_{2n}, $n \in \mathbf{Z}^+$

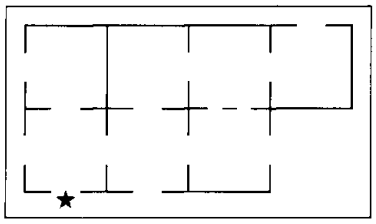


Figure 11.45

b) Find the maximum length of a circuit in

i) K_6 **ii)** K_8

iii) K_{10} **iv)** K_{2n}, $n \in \mathbf{Z}^+$

26. a) Let $G = (V, E)$ be a directed graph or multigraph with no isolated vertices. Prove that G has a directed Euler circuit if and only if G is connected and $od(v) = id(v)$ for all $v \in V$.

b) A directed graph is called *strongly connected* if there is a directed path from a to b for all vertices a, b, where $a \neq b$. Prove that if a directed graph has a directed Euler circuit, then it is strongly connected. Is the converse true?

27. Let G be a directed graph on n vertices. If the associated undirected graph for G is K_n, prove that $\sum_{v \in V}[od(v)]^2 = \sum_{v \in V}[id(v)]^2$.

28. If $G = (V, E)$ is a directed graph or multigraph with no isolated vertices, prove that G has a directed Euler trail if and only if (i) G is connected; (ii) $od(v) = id(v)$ for all but two vertices x, y in V; and (iii) $od(x) = id(x) + 1$, $id(y) = od(y) + 1$.

29. Let $V = \{000, 001, 010, \ldots, 110, 111\}$. For each four-bit sequence $b_1b_2b_3b_4$ draw an edge from the element $b_1b_2b_3$ to the element $b_2b_3b_4$ in V. (a) Draw the graph $G = (V, E)$ as described. (b) Find a directed Euler circuit for G. (c) Equally space eight 0's and eight 1's around the edge of a rotating (clockwise) drum so that these 16 bits form a circular sequence where the (consecutive) subsequences of length 4 provide the binary representations of 0, 1, 2, . . . , 14, 15 in some order.

30. Carolyn and Richard attended a party with three other married couples. At this party a good deal of handshaking took place, but (1) no one shook hands with her or his spouse; (2) no one shook hands with herself or himself; and (3) no one shook hands with anyone more than once. Before leaving the party, Carolyn asked the other seven people how many hands she or he had shaken. She received a different answer from each of the seven. How many times did Carolyn shake hands at this party? How many times did Richard?

31. Let $G = (V, E)$ be a loop-free connected undirected graph with $|V| \geq 2$. Prove that G contains two vertices v, w, where $\deg(v) = \deg(w)$.

32. If $G = (V, E)$ is an undirected graph with $|V| = n$ and $|E| = k$, the following matrices are used to represent G.

Let $V = \{v_1, v_2, \ldots, v_n\}$. Define the *adjacency matrix* $A = (a_{ij})_{n \times n}$ where $a_{ij} = 1$ if $\{v_i, v_j\} \in E$, otherwise $a_{ij} = 0$.

If $E = \{e_1, e_2, \ldots, e_k\}$, the *incidence matrix* I is the $n \times k$ matrix $(b_{ij})_{n \times k}$ where $b_{ij} = 1$ if v_i is a vertex on the edge e_j, otherwise $b_{ij} = 0$.

a) Find the adjacency and incidence matrices associated with the graph in Fig. 11.46.

b) Calculating A^2 and using the Boolean operations where $0 + 0 = 0, 0 + 1 = 1 + 0 = 1 + 1 = 1$, and $0 \cdot 0 = 0 \cdot 1 = 1 \cdot 0 = 0$, $1 \cdot 1 = 1$, prove that the entry in row i and column j of A^2 is 1 if and only if there is a walk of length 2 between the ith and jth vertices of V.

c) If we calculate A^2 using ordinary addition and multiplication, what do the entries in the matrix reveal about G?

d) What is the column sum for each column of A? Why?

e) What is the column sum for each column of I? Why?

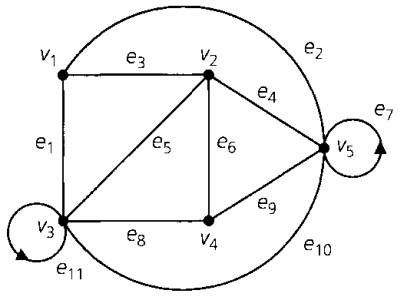


Figure 11.46

33. Determine whether or not the loop-free undirected graphs with the following adjacency matrices are isomorphic.

a) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

b) $\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$

c) $\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$

34. Determine whether or not the loop-free undirected graphs with the following incidence matrices are isomorphic.

a) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix}$

35. There are 15 people at a party. Is it possible for each of these people to shake hands with (exactly) three others?

36. Consider the two-by-four grid in Fig. 11.34. Assign the partial Gray code $A = \{00, 01, 11\}$ to the three horizontal levels: top (00), middle (01), and bottom (11). Now assign the partial Gray code $B = \{000, 001, 011, 010, 110\}$ to the five vertical levels: left, or first (000), second (001), third (011), fourth (010), and right, or fifth (110). Use the elements of $A \times B$ to label the 15 processors of this grid; for example, p_1 is labeled (00,000), p_2 is labeled (00, 001), p_8 is labeled (01, 011), p_{14} is labeled (11, 010), and p_{15} is labeled (11, 110). Show that the two-by-four grid is isomorphic to a subgraph of the hypercube Q_5. (Thus we can consider the two-by-four grid to be embedded in the hypercube Q_5.)

37. Prove that the three-by-three grid of Fig. 11.34 is isomorphic to a subgraph of the hypercube Q_4.

11.4 Planar Graphs

On a road map the lines indicating the roads and highways usually intersect only at junctions or towns. But sometimes roads seem to intersect when one road is located above another, as in the case of an overpass. In this case the two roads are at different levels, or planes. This type of situation leads us to the following definition.

Definition 11.17

A graph (or multigraph) G is called *planar* if G can be drawn in the plane with its edges intersecting only at vertices of G. Such a drawing of G is called an *embedding* of G in the plane.

EXAMPLE 11.15

The graphs in Fig. 11.47 are planar. The first is a 3-regular graph, because each vertex has degree 3; it is planar because no edges intersect except at the vertices. In graph (b) it appears that we have a *nonplanar* graph; the edges $\{x, z\}$ and $\{w, y\}$ overlap at a point other than a vertex. However, we can redraw this graph as shown in part (c) of the figure. Consequently, K_4 is planar.

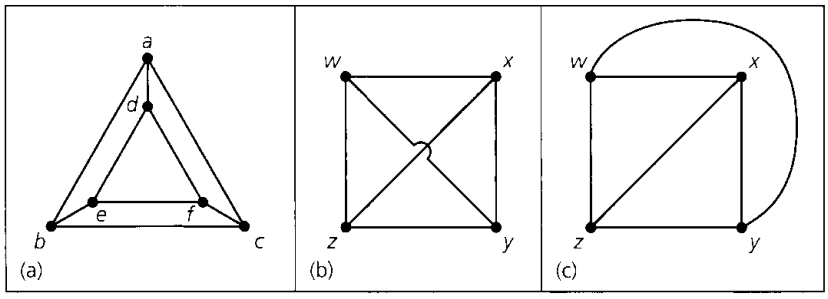


Figure 11.47

EXAMPLE 11.16

Just as K_4 is planar, so are the graphs K_1, K_2, and K_3.

An attempt to embed K_5 in the plane is shown in Fig. 11.48. If K_5 were planar, then any embedding would have to contain the pentagon in part (a) of the figure. Since a complete graph contains an edge for every pair of distinct vertices, we add edge $\{a, c\}$ as shown in part (b). This edge is contained entirely within the interior of the pentagon in part (a). (We could have drawn the edge in the exterior region determined by the pentagon. The reader will be asked in the exercises to show that the same conclusion arises in this case.) Moving

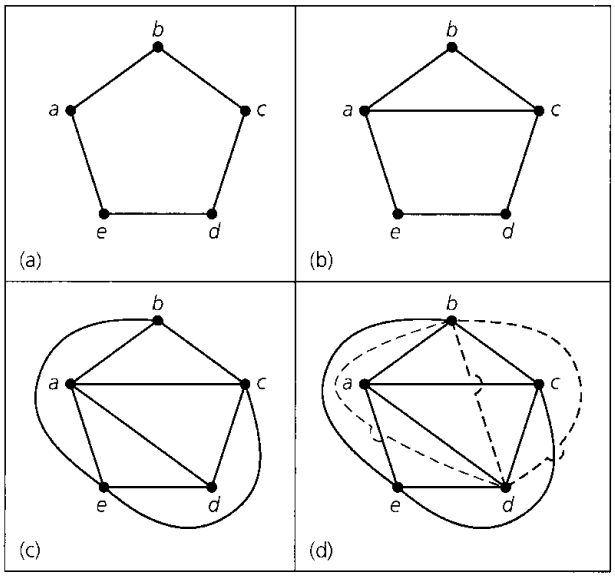


Figure 11.48

to part (c), we add in the edges $\{a, d\}$, $\{c, e\}$, and $\{b, e\}$. Now we consider the vertices b and d. We need the edge $\{b, d\}$ in order to have K_5. Vertex d is inside the region formed by the cycle edges $\{a, c\}$, $\{c, e\}$, and $\{e, a\}$, whereas b is outside the region. Thus in drawing the edge $\{b, d\}$, we must intersect one of the existing edges at least once, as shown by the dotted edges in part (d). Consequently, K_5 is nonplanar. (Since this proof appeals to a diagram, it definitely lacks rigor. However, later in the section we shall prove that K_5 is nonplanar by another method.)

Before we can characterize all nonplanar graphs we need to examine another class of graphs.

Definition 11.18

A graph $G = (V, E)$ is called *bipartite* if $V = V_1 \cup V_2$ with $V_1 \cap V_2 = \emptyset$, and every edge of G is of the form $\{a, b\}$ with $a \in V_1$ and $b \in V_2$. If each vertex in V_1 is joined with every vertex in V_2, we have a *complete bipartite* graph. In this case, if $|V_1| = m$, $|V_2| = n$, the graph is denoted by $K_{m,n}$.

EXAMPLE 11.17

Figure 11.49 indicates how we may partition the vertices of the hypercubes Q_1, Q_2, Q_3 to demonstrate that these graphs are bipartite. In general, for each $n \geq 1$, partition the vertices of Q_n as $V_1 \cup V_2$, where V_1 consists of all vertices whose binary labels have an even number of 1's, while V_2 consists of those whose binary labels have an odd number of 1's. Could there exist an edge $\{x, y\}$ in Q_n where $x, y \in V_1$? Recall that edges in Q_n connect vertices that differ in exactly one of the n positions in their binary labels. Suppose that the binary labels of x, y differ only in position i, for some $1 \leq i \leq n$. Then the total number of 1's in the binary labels for x, y is $2 \cdot$ [the number of 1's in x (or y) in all positions other than position i] $+ 1$, an odd total. But with $x, y \in V_1$, their binary labels each contain an even number of 1's — so the total number of 1's in these binary labels is even! This contradiction tells us that there is no edge $\{x, y\}$ in Q_n where $x, y \in V_1$. A similar argument can be given

to rule out the possibility of an edge $\{u, w\}$, where $u, w \in V_2$. Consequently, Q_n is bipartite for all $n \geq 1$.

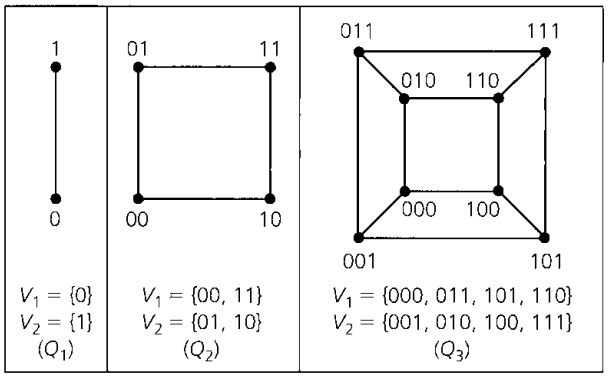


Figure 11.49

Figure 11.50 shows two bipartite graphs. The graph in part (a) satisfies the definition for $V_1 = \{a, b\}$ and $V_2 = \{c, d, e\}$. If we add the edges $\{b, d\}$ and $\{b, c\}$, the result is the complete bipartite graph $K_{2,3}$, which is planar. Graph (b) of the figure is $K_{3,3}$. Let $V_1 = \{h_1, h_2, h_3\}$ and $V_2 = \{u_1, u_2, u_3\}$, and interpret V_1 as a set of houses and V_2 as a set of utilities. Then $K_{3,3}$ is called the *utility graph*. Can we hook up each of the houses with each of the utilities and avoid having overlapping utility lines? In Fig. 11.50(b) it appears that this is not possible and that $K_{3,3}$ is nonplanar. (Once again we deduce the nonplanarity of a graph from a diagram. However, we shall verify that $K_{3,3}$ is nonplanar by another method, later in Example 11.21 of this section.)

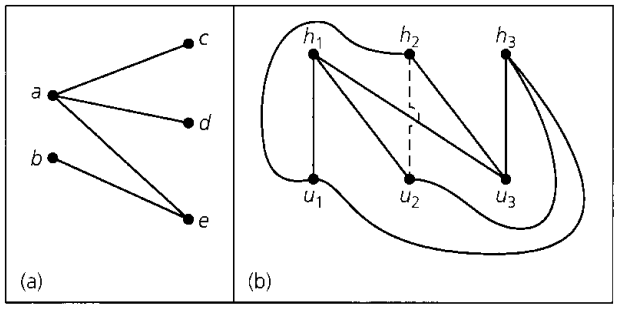


Figure 11.50

We shall see that when we are dealing with nonplanar graphs, either K_5 or $K_{3,3}$ will be the source of the problem. Before stating the general result, however, we need to develop one final new idea.

Definition 11.19

Let $G = (V, E)$ be a loop-free undirected graph, where $E \neq \emptyset$. An *elementary subdivision* of G results when an edge $e = \{u, w\}$ is removed from G and then the edges $\{u, v\}$, $\{v, w\}$ are added to $G - e$, where $v \notin V$.

The loop-free undirected graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are called *homeomorphic* if they are isomorphic or if they can both be obtained from the same loop-free undirected graph H by a sequence of elementary subdivisions.

EXAMPLE 11.18

a) Let $G = (V, E)$ be a loop-free undirected graph with $|E| \geq 1$. If G' is obtained from G by an elementary subdivision, then the graph $G' = (V', E')$ satisfies $|V'| = |V| + 1$ and $|E'| = |E| + 1$.

b) Consider the graphs G, G_1, G_2, and G_3 in Fig. 11.51. Here G_1 is obtained from G by means of one elementary subdivision: Delete edge $\{a, b\}$ from G and then add the edges $\{a, w\}$ and $\{w, b\}$. The graph G_2 is obtained from G by two elementary subdivisions. Hence G_1 and G_2 are homeomorphic. Also, G_3 can be obtained from G by four elementary subdivisions, so G_3 is homeomorphic to both G_1 and G_2.

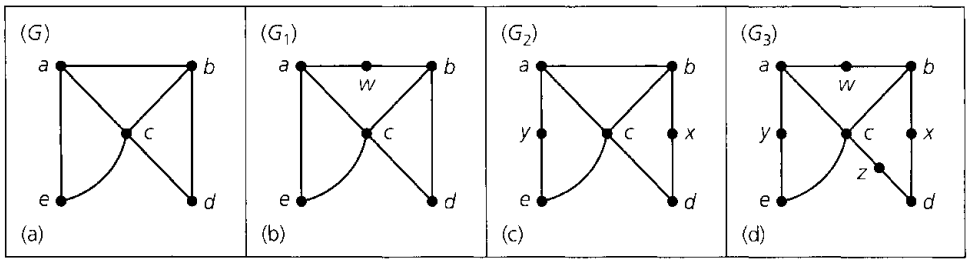


Figure 11.51

However, we cannot obtain G_1 from G_2 (or G_2 from G_1) by a sequence of elementary subdivisions. Furthermore, the graph G_3 can be obtained from either G_1 or G_2 by a sequence of elementary subdivisions: six (such sequences of three elementary subdivisions) for G_1 and two for G_2. But neither G_1 nor G_2 can be obtained from G_3 by a sequence of elementary subdivisions.

One may think of homeomorphic graphs as being isomorphic except, possibly, for vertices of degree 2. In particular, if two graphs are homeomorphic, they are either both planar or they are both nonplanar.

These preliminaries lead us to the following result.

THEOREM 11.5

Kuratowski's Theorem. A graph is nonplanar if and only if it contains a subgraph that is homeomorphic to either K_5 or $K_{3,3}$.

Proof: (This theorem was first proved by the Polish mathematician Kasimir Kuratowski in 1930.) If a graph G has a subgraph homeomorphic to either K_5 or $K_{3,3}$, it is clear that G is nonplanar. The converse of this theorem, however, is much more difficult to prove. (A proof can be found in Chapter 8 of C. L. Liu [23] or Chapter 6 of D. B. West [32].)

We demonstrate the use of Kuratowski's Theorem in the following example.

EXAMPLE 11.19

a) Figure 11.52(a) is a familiar graph called the *Petersen graph.* Part (b) of the figure provides a subgraph of the Petersen graph that is homeomorphic to $K_{3,3}$. (Figure 11.53 shows how the subgraph is obtained from $K_{3,3}$ by a sequence of four elementary subdivisions.) Hence the Petersen graph is nonplanar.

b) In part (a) of Fig. 11.54 we find the 3-regular graph G, which is isomorphic to the 3-dimensional hypercube Q_3. The 4-regular complement of G is shown in Fig. 11.54(b), where the edges $\{a, g\}$ and $\{d, f\}$ suggest that G *may be* nonplanar. Figure 11.54(c)

depicts a subgraph H of $\overline{G}$ that is homeomorphic to K_5, so by Kuratowski's Theorem it follows that $\overline{G}$ *is* nonplanar.

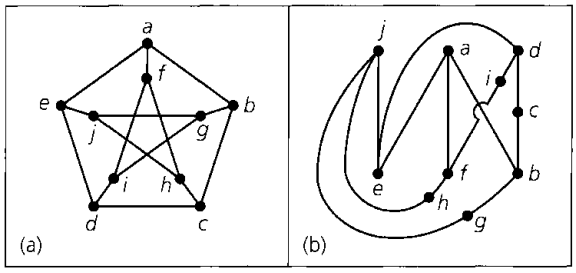


Figure 11.52

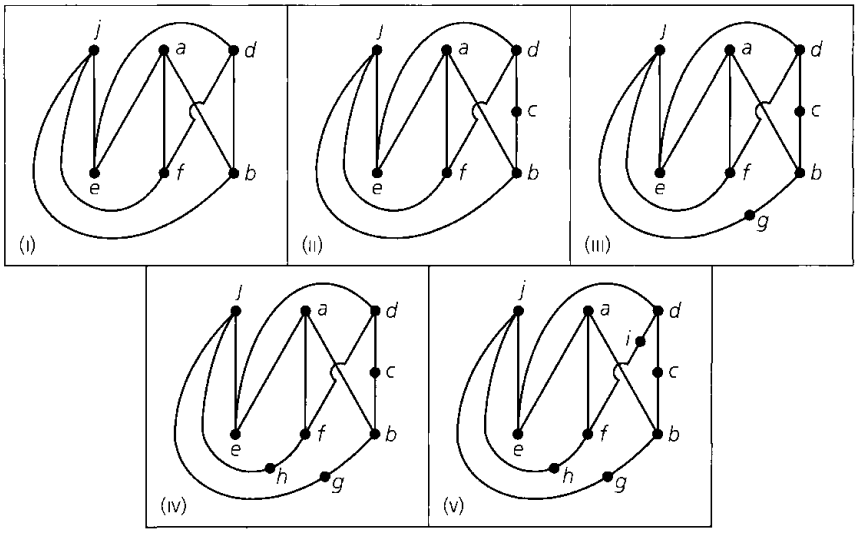


Figure 11.53

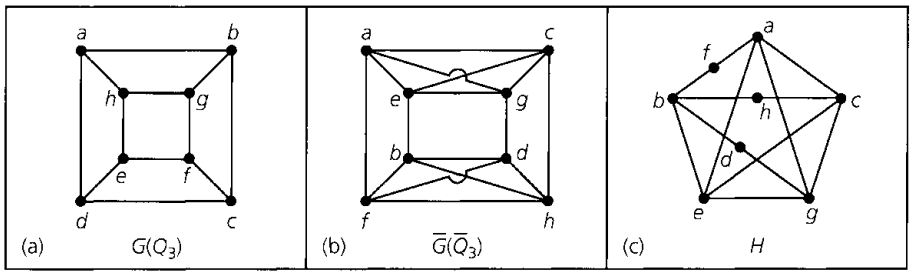


Figure 11.54

When a graph or multigraph is planar and connected, we find the following relation, which was discovered by Euler. For this relation we need to be able to count the number of regions determined by a planar connected graph or multigraph — the number (of these regions) being defined only when we have a planar embedding of the graph. For instance, the planar embedding of K_4 in part (a) of Fig. 11.55 demonstrates how this depiction of K_4 determines four regions in the plane: three of finite area — namely, R_1, R_2, and R_3 — and

the infinite region R_4. When we look at Fig. 11.55(b) we might think that here K_4 determines five regions, but this depiction does *not* present a planar embedding of K_4. So the result in Fig. 11.55(a) is the only one we actually want to deal with here.

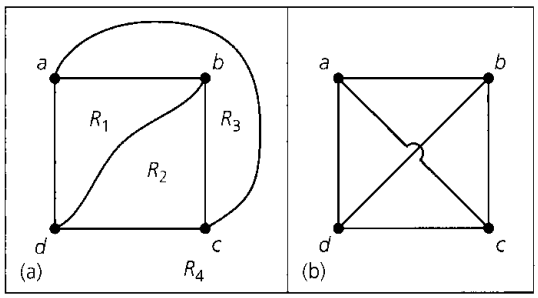


Figure 11.55

THEOREM 11.6

Let $G = (V, E)$ be a connected planar graph or multigraph with $|V| = v$ and $|E| = e$. Let r be the number of regions in the plane determined by a planar embedding (or, depiction) of G; one of these regions has infinite area and is called *the infinite region*. Then $v - e + r = 2$.

Proof: The proof is by induction on e. If $e = 0$ or 1, then G is isomorphic to one of the graphs in Fig. 11.56. The graph in part (a) has $v = 1, e = 0$, and $r = 1$; so, $v - e + r = 1 - 0 + 1 = 2$. For graph (b), $v = 1, e = 1$, and $r = 2$. Graph (c) has $v = 2, e = 1$, and $r = 1$. In both cases, $v - e + r = 2$.

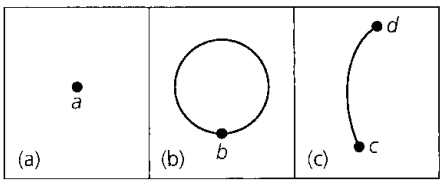


Figure 11.56

Now let $k \in \mathbf{N}$ and assume that the result is true for every connected planar graph or multigraph with e edges, where $0 \leq e \leq k$. If $G = (V, E)$ is a connected planar graph or multigraph with v vertices, r regions, and $e = k + 1$ edges, let $a, b \in V$ with $\{a, b\} \in E$. Consider the subgraph H of G obtained by deleting the edge $\{a, b\}$ from G. (If G is a multigraph and $\{a, b\}$ is one of a set of edges between a and b, then we remove it only once.) Consequently, we may write $H = G - \{a, b\}$ or $G = H + \{a, b\}$. We consider the following two cases, depending on whether H is connected or disconnected.

Case 1: The results in parts (a), (b), (c), and (d) of Fig. 11.57 show us how a graph G may be obtained from a connected graph H when the (new) loop $\{a, a\}$ is drawn as in parts (a) and (b) or when the (new) edge $\{a, b\}$ joins two distinct vertices in H as in parts (c) and (d). In all of these situations, H has v vertices, k edges, and $r - 1$ regions because one of the regions for H is split into two regions for G. The induction hypothesis applied to graph H tells us that $v - k + (r - 1) = 2$, and from this it follows that $2 = v - (k + 1) + r = v - e + r$. So Euler's Theorem is true for G in this case.

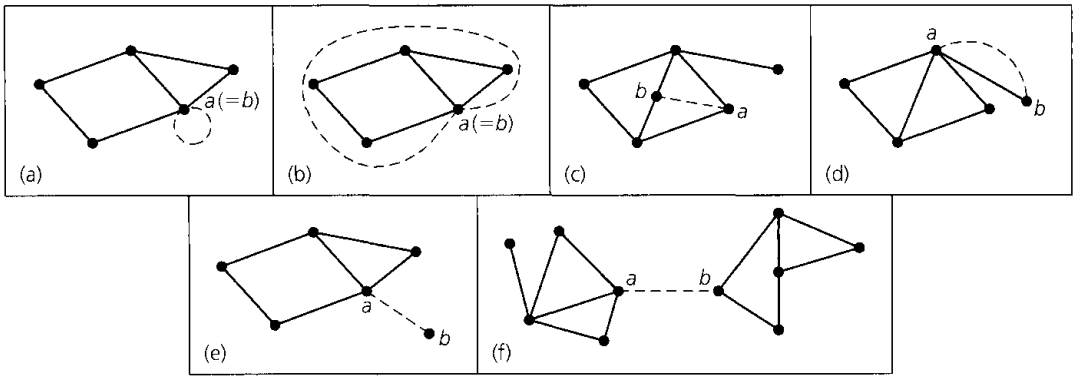


Figure 11.57

Case 2: Now we consider the case where $G - \{a, b\} = H$ is a disconnected graph [as demonstrated in Fig. 11.57(e) and (f)]. Here H has v vertices, k edges, and r regions. Also, H has two components H_1 and H_2, where H_i has v_i vertices, e_i edges, and r_i regions, for $i = 1, 2$. [Part (e) of Fig. 11.57 indicates that one component could consist of just an isolated vertex.] Furthermore, $v_1 + v_2 = v$, $e_1 + e_2 = k\ (= e - 1)$, and $r_1 + r_2 = r + 1$ because each of H_1 and H_2 determines an infinite region. When we apply the induction hypothesis to each of H_1 and H_2 we learn that

$$v_1 - e_1 + r_1 = 2 \quad \text{and} \quad v_2 - e_2 + r_2 = 2.$$

Consequently, $(v_1 + v_2) - (e_1 + e_2) + (r_1 + r_2) = v - (e - 1) + (r + 1) = 4$, and from this it follows that $v - e + r = 2$, thus establishing Euler's Theorem for G in this case.

The following corollary for Theorem 11.6 provides two inequalities relating the number of edges in a loop-free connected planar graph G with (1) the number of regions determined by a planar embedding of G; and (2) the number of vertices in G. Before we examine this corollary, however, let us look at the following helpful idea. For each region R in a planar embedding of a (planar) graph or multigraph, the *degree of* R, denoted $\deg(R)$, is the number of edges traversed in a (shortest) closed walk about (the edges in) the boundary of R. If $G = (V, E)$ is the graph of Fig. 11.58(a), then this planar embedding of G has four regions where

$$\deg(R_1) = 5, \qquad \deg(R_2) = 3, \qquad \deg(R_3) = 3, \qquad \deg(R_4) = 7.$$

[Here $\deg(R_4) = 7$, as determined by the closed walk: $a \to b \to g \to h \to g \to f \to d \to a$.] Part (b) of the figure shows a second planar embedding of G — again with four regions — and here

$$\deg(R_5) = 4, \qquad \deg(R_6) = 3, \qquad \deg(R_7) = 5, \qquad \deg(R_8) = 6.$$

[The closed walk $b \to g \to h \to g \to f \to b$ gives us $\deg(R_7) = 5$.)]

We see that $\sum_{i=1}^{4} \deg(R_i) = 18 = \sum_{i=5}^{8} \deg(R_i) = 2 \cdot 9 = 2|E|$. This is true in general because each edge of the planar embedding is either part of the boundary of two regions [like $\{b, c\}$ in parts (a) and (b)] or occurs twice in the closed walk about the edges in the boundary for one region [like $\{g, h\}$ in parts (a) and (b)].

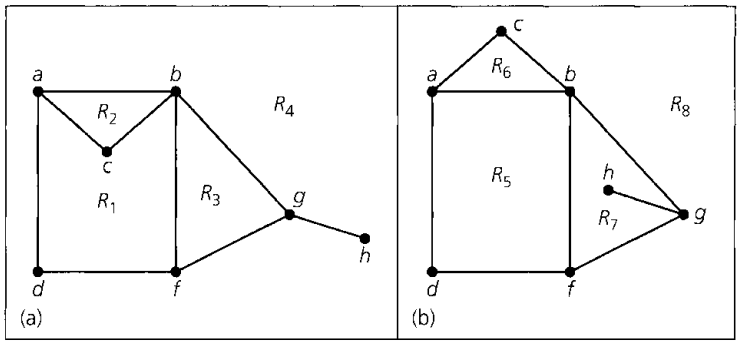


Figure 11.58

Now let us consider the following.

COROLLARY 11.3

Let $G = (V, E)$ be a loop-free connected planar graph with $|V| = v$, $|E| = e > 2$, and r regions. Then $3r \leq 2e$ and $e \leq 3v - 6$.

Proof: Since G is loop-free and is not a multigraph, the boundary of each region (including the infinite region) contains at least three edges—hence, each region has degree ≥ 3. Consequently, $2e = 2|E|$ = the sum of the degrees of the r regions determined by G and $2e \geq 3r$. From Euler's Theorem, $2 = v - e + r \leq v - e + (2/3)e = v - (1/3)e$, so $6 \leq 3v - e$, or $e \leq 3v - 6$.

We now consider what this corollary does and does not imply. If $G = (V, E)$ is a loop-free connected graph with $|E| > 2$, then if $e > 3v - 6$, it follows that G is not planar. However, if $e \leq 3v - 6$, we cannot conclude that G is planar.

EXAMPLE 11.20

The graph K_5 is loop-free and connected with ten edges and five vertices. Consequently, $3v - 6 = 15 - 6 = 9 < 10 = e$. Therefore, by Corollary 11.3, we find that K_5 is nonplanar.

EXAMPLE 11.21

The graph $K_{3,3}$ is loop-free and connected with nine edges and six vertices. Here $3v - 6 = 18 - 6 = 12 \geq 9 = e$. It would be a mistake to conclude from this that $K_{3,3}$ is planar. It would be the mistake of arguing by the converse.

However, $K_{3,3}$ is nonplanar. If $K_{3,3}$ were planar, then since each region in the graph is bounded by at least four edges, we have $4r \leq 2e$. (We found a similar situation in the proof of Corollary 11.3.) From Euler's Theorem, $v - e + r = 2$, or $r = e - v + 2 = 9 - 6 + 2 = 5$, so $20 = 4r \leq 2e = 18$. From this contradiction we have $K_{3,3}$ being nonplanar.

EXAMPLE 11.22

We use Euler's Theorem to characterize the *Platonic solids*. [For these solids all faces are congruent and all (interior) solid angles are equal.] In Fig. 11.59 we have two of these solids. Part (a) of the figure shows the regular tetrahedron, which has four faces, each an equilateral triangle. Concentrating on the edges of the tetrahedron, we focus on its underlying framework. As we view this framework from a point directly above the center of one of the faces, we picture the planar representation in part (b). This planar graph determines four regions (corresponding to the four faces); three regions meet at each of the four vertices. Part (c) of the figure provides another Platonic solid, the cube. Its associated planar graph is given in part (d). In this graph there are six regions with three regions meeting at each vertex.

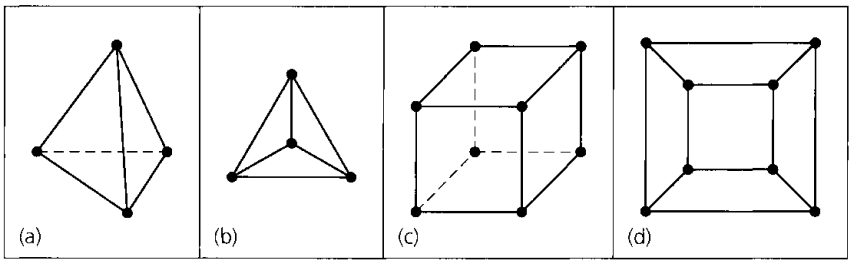


Figure 11.59

On the basis of our observations for the regular tetrahedron and the cube, we shall determine the other Platonic solids by means of their associated planar graphs. In these graphs $G = (V, E)$ we have $v = |V|$; $e = |E|$; r = the number of planar regions determined by G; m = the number of edges in the boundary of each region; and n = the number of regions that meet at each vertex. Thus the constants $m, n \geq 3$. Since each edge is used in the boundary of two regions and there are r regions, each with m edges, it follows that $2e = mr$. Counting the endpoints of the edges, we get $2e$. But all these endpoints can also be counted by considering what happens at each vertex. Since n regions meet at each vertex, n edges meet there, so there are n endpoints of edges to count at each of the v vertices. This totals nv endpoints of edges, so $2e = nv$. From Euler's Theorem we have

$$0 < 2 = v - e + r = \frac{2e}{n} - e + \frac{2e}{m} = e\left(\frac{2m - mn + 2n}{mn}\right).$$

With $e, m, n > 0$, we find that

$$2m - mn + 2n > 0 \Rightarrow mn - 2m - 2n < 0$$
$$\Rightarrow mn - 2m - 2n + 4 < 4 \Rightarrow (m-2)(n-2) < 4.$$

Since $m, n \geq 3$, we have $(m-2), (n-2) \in \mathbf{Z}^+$, and there are only five cases to consider:

1) $(m-2) = (n-2) = 1$; $m = n = 3$ (The regular tetrahedron)

2) $(m-2) = 2, (n-2) = 1$; $m = 4, n = 3$ (The cube)

3) $(m-2) = 1, (n-2) = 2$; $m = 3, n = 4$ (The octahedron)

4) $(m-2) = 3, (n-2) = 1$; $m = 5, n = 3$ (The dodecahedron)

5) $(m-2) = 1, (n-2) = 3$; $m = 3, n = 5$ (The icosahedron)

The planar graphs for cases 3–5 are shown in Fig. 11.60.

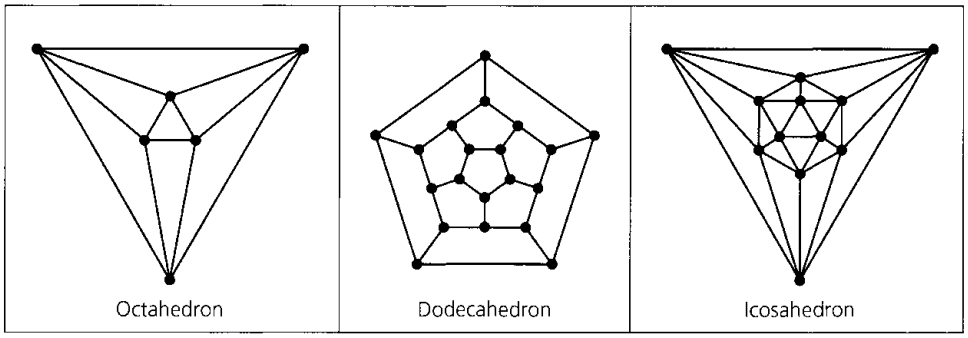


Figure 11.60

The last idea we shall discuss for planar graphs is the notion of a *dual* graph. This concept is also valid for planar graphs with loops and for planar multigraphs. To construct a dual (relative to a particular embedding) for a planar graph or multigraph G with $V = \{a, b, c, d, e, f\}$, place a point (vertex) inside each region, including the infinite region, determined by the graph, as in Fig. 11.61(a). For each edge shared by two regions, draw an edge connecting the vertices inside these regions. For an edge that is traversed twice in the closed walk about the edges of one region, draw a loop at the vertex for this region. In Fig. 11.61(b), G^d is a dual for the graph $G = (V, E)$. From this example we make the following observations:

1) An edge in G corresponds with an edge in G^d, and conversely.
2) A vertex of degree 2 in G yields a pair of edges in G^d that connect the same two vertices. Hence G^d may be a multigraph. (Here vertex e provides the edges $\{a, e\}$, $\{e, f\}$ in G that brought about the two edges connecting v and z in G^d.)
3) Given a loop in G, if the interior of the (finite area) region determined by the loop contains no other vertex or edge of G, then the loop yields a pendant vertex in G^d. (It is also true that a pendant vertex in G yields a loop in G^d.)
4) The degree of a vertex in G^d is the number of edges in the boundary of the closed walk about the region in G that contains that vertex.

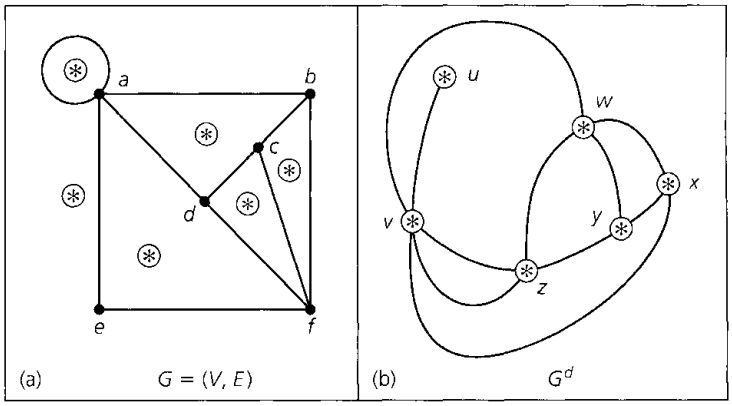


Figure 11.61

(Why is G^d called *a* dual of G instead of *the* dual of G? The Section Exercises will show that it is possible to have isomorphic graphs G_1 and G_2 with respective duals G_1^d, G_2^d that are not isomorphic.)

In order to examine further the relationship between a graph G and a dual G^d of G, we introduce the following idea. [Here we recall (from Definition 11.5) that $\kappa(G)$ counts the number of components of G.]

Definition 11.20 Let $G = (V, E)$ be an undirected graph or multigraph. A subset E' of E is called a *cut-set* of G if by removing the edges (but not the vertices) in E' from G, we have $\kappa(G) < \kappa(G')$, where $G' = (V, E - E')$; but when we remove (from E) any proper subset E'' of E', we have $\kappa(G) = \kappa(G'')$, for $G'' = (V, E - E'')$.

EXAMPLE 11.23

For a given connected graph, a cut-set is a *minimal* disconnecting set of edges. In the graph in Fig. 11.62(a), note that each of the sets $\{\{a, b\}, \{a, c\}\}$, $\{\{a, b\}, \{c, d\}\}$, $\{\{e, h\}, \{f, h\}, \{g, h\}\}$, and $\{\{d, f\}\}$ is a cut-set. For the graph in part (b) of the figure, the edge set $\{\{n, p\}, \{r, p\}, \{r, s\}\}$ is a cut-set. Note that the edges in this cut-set are *not* all incident to some single vertex. Here the cut-set separates the vertices m, n, r from the vertices p, s, t. The edge set $\{\{s, t\}\}$ is also a cut-set for this graph — the removal of the edge $\{s, t\}$ from this connected graph results in a subgraph with two components, one of which is the isolated vertex t.

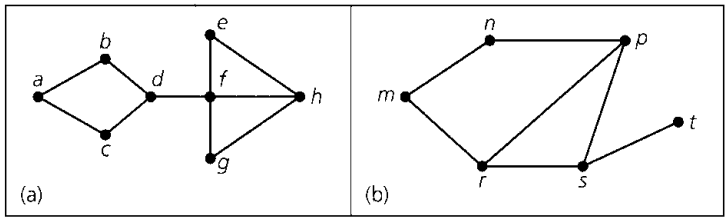


Figure 11.62

Whenever a cut-set for a connected graph consists of only one edge, that edge is called a *bridge* for the graph. For the graph in Fig. 11.62(a), the edge $\{d, f\}$ is the only bridge; the edge $\{s, t\}$ is the only bridge in part (b) of the figure.

We return now to the graphs in Fig. 11.61, redrawing them as shown in Fig. 11.63 in order to emphasize the correspondence between their edges.

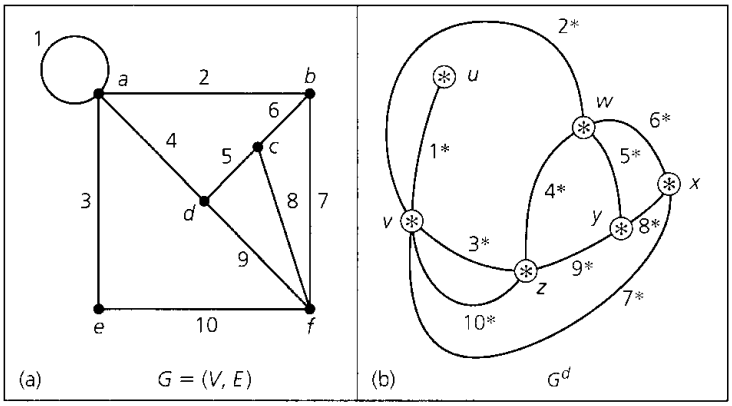


Figure 11.63

Here the edges in G are labeled $1, 2, \ldots, 10$. The numbering scheme for G^d is obtained as follows: The edge labeled 4*, for example, connects the vertices w and z in G^d. We drew this edge because edge 4 in G was a common edge of the regions containing these vertices. Likewise, edge 7 is common to the region containing x and the infinite region containing v. Hence we label the edge in G^d that connects x and v with 7*.

In graph G the set of edges labeled 6, 7, 8 constitutes a cycle. What about the edges labeled 6*, 7*, 8* in G^d? If they are removed from G^d, then vertex x becomes isolated and G^d is disconnected. Since we cannot disconnect G^d by removing any proper subset

of {6*, 7*, 8*}, these edges form a cut-set in G^d. In similar fashion, edges 2, 4, 10 form a cut-set in G, whereas in G^d the edges 2*, 4*, 10* yield a cycle.

We also have the two-edge cut-set {3, 10} in G, and we find that the edges 3*, 10* provide a two-edge circuit in G^d. Another observation: The one-edge cut set {1*} in G^d comes about from edge 1, a loop in G.

In general, there is a one-to-one correspondence between the following sets of edges in a planar graph G and a dual G^d of G.

1) Cycles (cut-sets) of n (≥ 3) edges in G correspond with cut-sets (cycles) of n edges in G^d.
2) A loop in G corresponds with a one-edge cut-set in G^d.
3) A one-edge cut-set in G corresponds with a loop in G^d.
4) A two-edge cut-set in G corresponds with a two-edge circuit in G^d.
5) If G is a planar multigraph, then each two-edge circuit in G determines a two-edge cut-set of G^d.

All these theoretical observations are interesting, but let us stop here and see how we might apply the idea of a dual.

EXAMPLE 11.24

If we consider the five finite regions in Fig. 11.64(a) as countries on a map, and we construct the subgraph (because we do not use the infinite region) of a dual as shown in part (b), then we find the following relationship.

Suppose we are confronted with the "mapmaker's problem" whereby we want to color the five regions of the map in part (a) so that two countries that share a common border are colored with different colors. This type of coloring can be translated into the dual notion of coloring the vertices in part (b) so that adjacent vertices are colored with different colors. (Such coloring problems will be examined further in Section 11.6.)

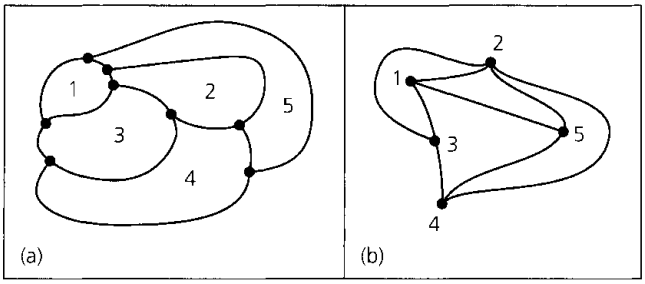


Figure 11.64

The final result for this section provides us with an application for an electrical network. This material is based on Example 8.6 on pp. 227–230 of the text by C. L. Liu [23].

EXAMPLE 11.25

In Fig. 11.65 we see an electrical network with nine contacts (switches) that control the excitation of a light. We want to construct a dual network where a second light will go on (off) whenever the light in our given network is off (on).

The contacts (switches) are of two types: normally open (as shown in Fig. 11.65) and normally closed. We use a and a' as in Fig. 11.66 to represent the normally open and normally closed contacts, respectively.

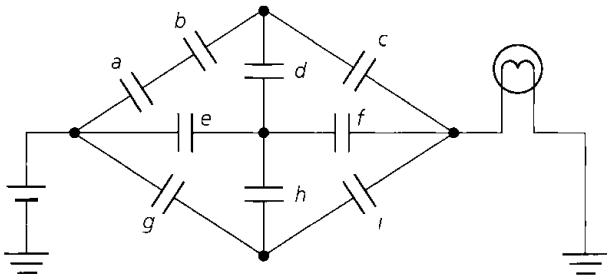


Figure 11.65

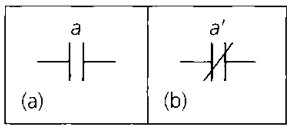


Figure 11.66

In Fig. 11.67(a) a *one-terminal-pair-graph* represents the network in Fig. 11.65. Here the special pair of vertices is labeled 1 and 2. These vertices are called the *terminals* of the graph. Also each edge is labeled according to its corresponding contact in Fig. 11.65.

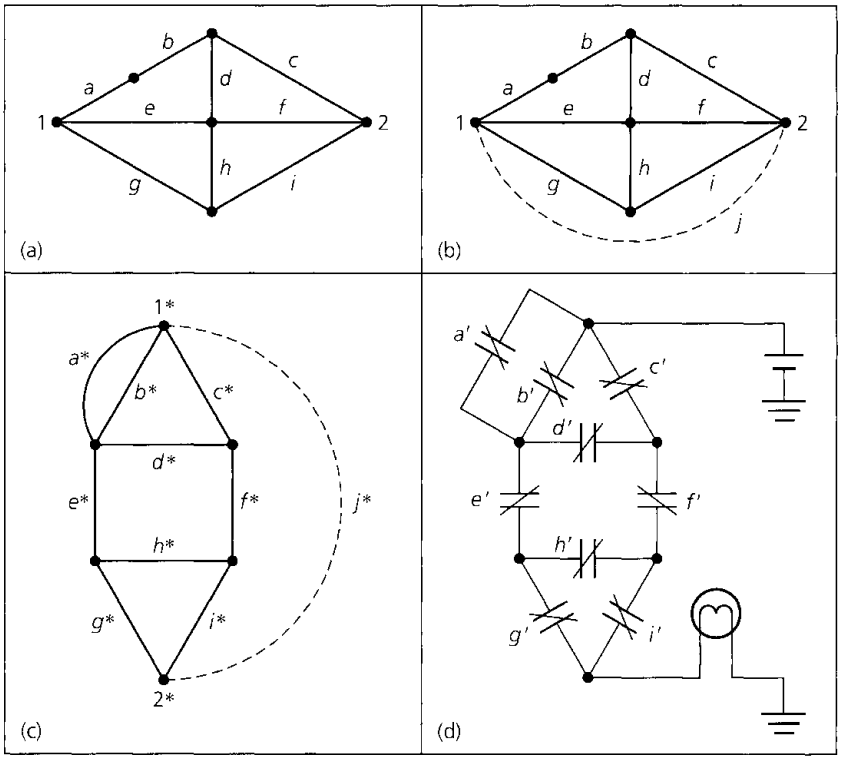


Figure 11.67

A one-terminal-pair-graph G is called a *planar-one-terminal-pair-graph* if G is planar, and the resulting graph is also planar when an edge connecting the terminals is added to G. Figure 11.67(b) shows this situation. Constructing a dual of part (b), we obtain the graph in part (c) of the figure. Removal of the dotted edge results in the terminals 1*, 2* for this dual, which is a one-terminal-pair-graph. This graph provides the dual network in Fig. 11.67(d).

We make two observations in closing.

1) When the contacts at a, b, c are closed in the original network (Fig. 11.65), the light is on. In Fig. 11.67(b) the edges a, b, c, j form a cycle that includes the terminals.

In part (c) of the figure, the edges a^*, b^*, c^*, j^* form a cut-set disconnecting the terminals 1*, 2*. Finally, with a', b', c' open in part (d) of the figure, no current gets past the first level of contacts (switches) and the light is off.

2) In like manner, the edges c, d, e, g, j form a cut-set that separates the terminals in Fig. 11.67(b). (When the contacts at c, d, e, g are open in Fig. 11.65, the light is off.) Figure 11.67(c) shows how c^*, d^*, e^*, g^*, j^* form a cycle that includes 1*, 2*. If c', d', e', g' are closed in part (d), current flows through the dual network and the light is on.

EXERCISES 11.4

1. Verify that the conclusion in Example 11.16 is unchanged if Fig. 11.48(b) has edge $\{a, c\}$ drawn in the exterior of the pentagon.

2. Show that when any edge is removed from K_5, the resulting subgraph is planar. Is this true for the graph $K_{3,3}$?

3. **a)** How many vertices and how many edges are there in the complete bipartite graphs $K_{4,7}$, $K_{7,11}$, and $K_{m,n}$, where $m, n, \in \mathbf{Z}^+$?

b) If the graph $K_{m,12}$ has 72 edges, what is m?

4. Prove that any subgraph of a bipartite graph is bipartite.

5. For each graph in Fig. 11.68 determine whether or not the graph is bipartite.

6. Let $n \in \mathbf{Z}^+$ with $n \geq 4$. How many subgraphs of K_n are isomorphic to the complete bipartite graph $K_{1,3}$?

7. Let $m, n \in \mathbf{Z}^+$ with $m \geq n \geq 2$. (a) Determine how many distinct cycles of length 4 there are in $K_{m,n}$. (b) How many different paths of length 2 are there in $K_{m,n}$? (c) How many different paths of length 3 are there in $K_{m,n}$?

8. What is the length of a longest path in each of the following graphs?

a) $K_{1,4}$ **b)** $K_{3,7}$ **c)** $K_{7,12}$

d) $K_{m,n}$, where $m, n \in \mathbf{Z}^+$ with $m < n$.

9. How many paths of longest length are there in each of the following graphs? (Remember that a path such as $v_1 \to v_2 \to v_3$ is considered to be the same as the path $v_3 \to v_2 \to v_1$.)

a) $K_{1,4}$ **b)** $K_{3,7}$ **c)** $K_{7,12}$

d) $K_{m,n}$, where $m, n \in \mathbf{Z}^+$ with $m < n$.

10. Can a bipartite graph contain a cycle of odd length? Explain.

11. Let $G = (V, E)$ be a loop-free connected graph with $|V| = v$. If $|E| > (v/2)^2$, prove that G cannot be bipartite.

12. **a)** Find all the nonisomorphic complete bipartite graphs $G = (V, E)$, where $|V| = 6$.

b) How many nonisomorphic complete bipartite graphs $G = (V, E)$ satisfy $|V| = n \geq 2$?

13. **a)** Let $X = \{1, 2, 3, 4, 5\}$. Construct the loop-free undirected graph $G = (V, E)$ as follows:

- (V): Let each two-element subset of X represent a vertex in G.
- (E): If $v_1, v_2 \in V$ correspond to subsets $\{a, b\}$ and $\{c, d\}$, respectively, of X, then draw the edge $\{v_1, v_2\}$ in G if $\{a, b\} \cap \{c, d\} = \emptyset$.

b) To what graph is G isomorphic?

14. Determine which of the graphs in Fig. 11.69 are planar. If a graph is planar, redraw it with no edges overlapping. If it is nonplanar, find a subgraph homeomorphic to either K_5 or $K_{3,3}$.

15. Let $m, n \in \mathbf{Z}^+$ with $m \leq n$. Under what condition(s) on m, n will every edge in $K_{m,n}$ be in exactly one of two isomorphic subgraphs of $K_{m,n}$?

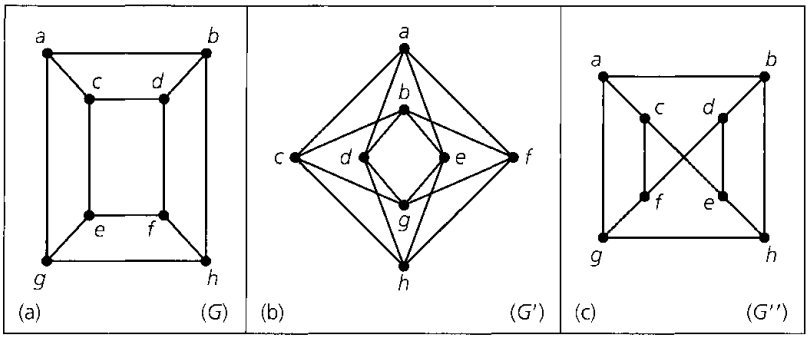


Figure 11.68

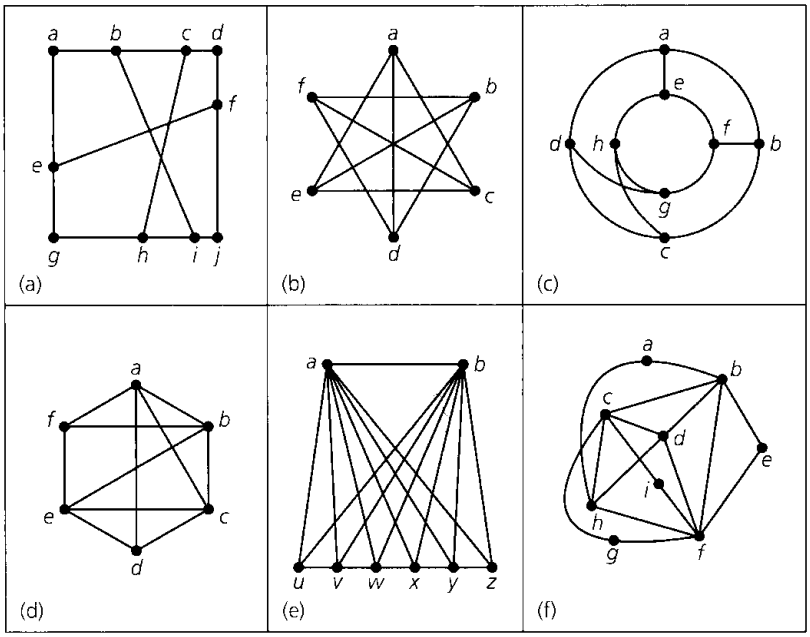


Figure 11.69

16. Prove that the Petersen graph is isomorphic to the graph in Fig. 11.70

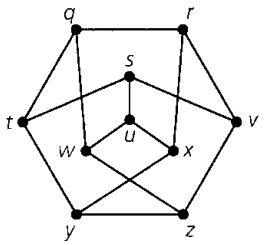


Figure 11.70

17. Determine the number of vertices, the number of edges, and the number of regions for each of the planar graphs in Fig. 11.71. Then show that your answers satisfy Euler's Theorem for connected planar graphs.

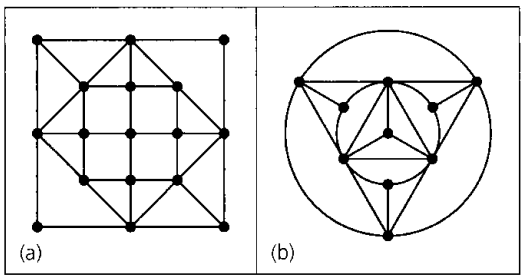


Figure 11.71

18. Let $G = (V, E)$ be an undirected connected loop-free graph. Suppose further that G is planar and determines 53 regions. If, for some planar embedding of G, each region has at least five edges in its boundary, prove that $|V| \geq 82$.

19. Let $G = (V, E)$ be a loop-free connected 4-regular planar graph. If $|E| = 16$, how many regions are there in a planar depiction of G?

20. Suppose that $G = (V, E)$ is a loop-free planar graph with $|V| = v$, $|E| = e$, and $\kappa(G) =$ the number of components of G. (a) State and prove an extension of Euler's Theorem for such a graph. (b) Prove that Corollary 11.3 remains valid if G is loop-free and planar but not connected.

21. Prove that every loop-free connected planar graph has a vertex v with $\deg(v) < 6$.

22. **a)** Let $G = (V, E)$ be a loop-free connected graph with $|V| \geq 11$. Prove that either G or its complement $\overline{G}$ must be nonplanar.

b) The result in part (a) is actually true for $|V| \geq 9$, but the proof for $|V| = 9, 10$, is much harder. Find a counterexample to part (a) for $|V| = 8$.

23. **a)** Let $k \in \mathbf{Z}^+$, $k \geq 3$. If $G = (V, E)$ is a connected planar graph with $|V| = v$, $|E| = e$, and each cycle of length at least k, prove that $e \leq \left(\frac{k}{k-2}\right)(v-2)$.

b) What is the minimal cycle length in $K_{3,3}$?

c) Use parts (a) and (b) to conclude that $K_{3,3}$ is nonplanar.

d) Use part (a) to prove that the Petersen graph is nonplanar.

24. **a)** Find a dual graph for each of the two planar graphs and the one planar multigraph in Fig. 11.72.

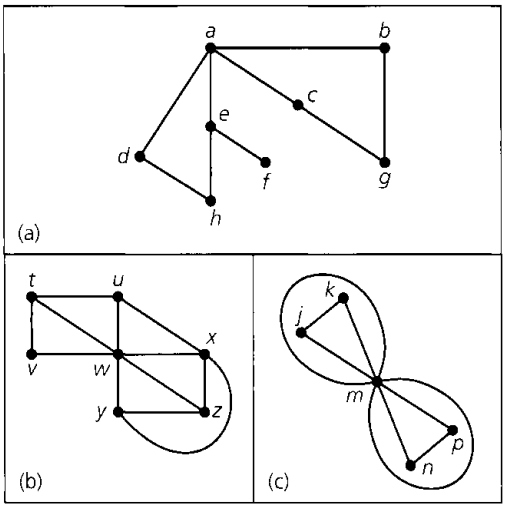


Figure 11.72

b) Does the dual for the multigraph in part (c) have any pendant vertices? If not, does this contradict the third observation made prior to Definition 11.20?

25. a) Find duals for the planar graphs that correspond with the five Platonic solids.

b) Find the dual of the graph W_n, the wheel with n spokes (as defined in Exercise 14 of Section 11.1).

26. a) Show that the graphs in Fig. 11.73 are isomorphic.

b) Draw a dual for each graph.

c) Show that the duals obtained in part (b) are not isomorphic.

d) Two graphs G and H are called *2-isomorphic* if one can be obtained from the other by applying either or both of the following procedures a finite number of times.

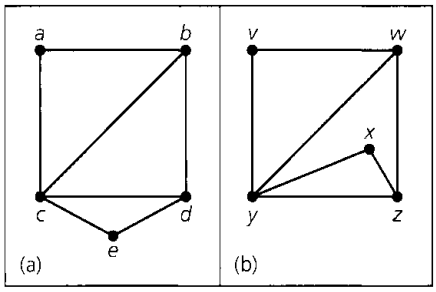


Figure 11.73

1) In Fig. 11.74 we split a vertex, namely r, of G and obtain the graph H, which is disconnected.

2) In Fig. 11.75 we obtain graph (d) from graph (a) by

- **i)** first splitting the two distinct vertices j and q — disconnecting the graph,
- **ii)** then reflecting one subgraph about the horizontal axis, and
- **iii)** then identifying vertex $j(q)$ in one subgraph with vertex $q(j)$ in the other subgraph.

Prove that the dual graphs obtained in part (c) are 2-isomorphic.

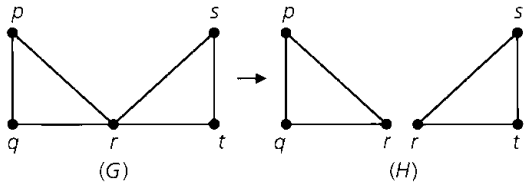


Figure 11.74

e) For the cut-set $\{\{a, b\}, \{c, b\}, \{d, b\}\}$ in part (a) of Fig. 11.73, find the corresponding cycle in its dual. In the

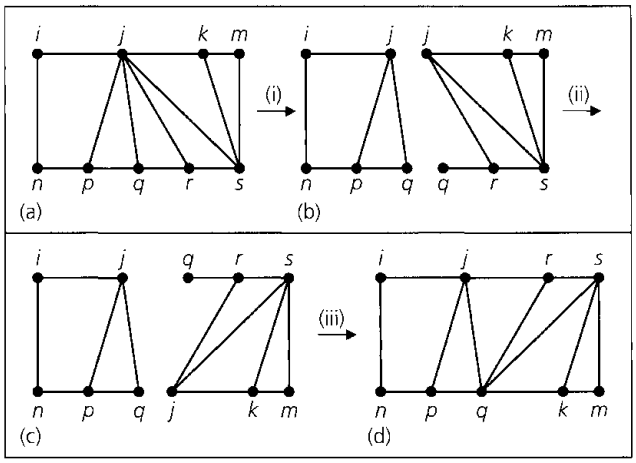


Figure 11.75

dual of the graph in Fig. 11.73(b), find the cut-set that corresponds with the cycle $\{w, z\}, \{z, x\}, \{x, y\}, \{y, w\}$ in the given graph.

27. Find the dual network for the electrical network shown in Fig. 11.76.

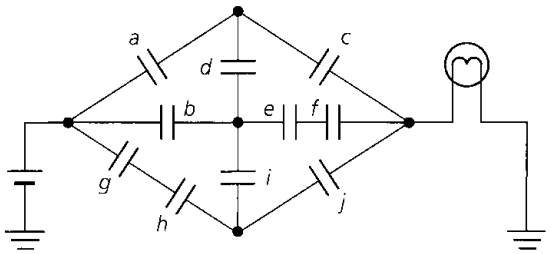


Figure 11.76

28. Let $G = (V, E)$ be a loop-free connected planar graph. If G is isomorphic to its dual and $|V| = n$, what is $|E|$?

29. Let G_1, G_2 be two loop-free connected undirected graphs. If G_1, G_2 are homeomorphic, prove that (a) G_1, G_2 have the same number of vertices of odd degree; (b) G_1 has an Euler trail if and only if G_2 has an Euler trail; and (c) G_1 has an Euler circuit if and only if G_2 has an Euler circuit.

11.5 Hamilton Paths and Cycles

In 1859 the Irish mathematician Sir William Rowan Hamilton (1805–1865) developed a game that he sold to a Dublin toy manufacturer. The game consisted of a wooden regular dodecahedron with the 20 corner points (vertices) labeled with the names of prominent cities. The objective of the game was to find a cycle along the edges of the solid so that each city was on the cycle (exactly once). Figure 11.77 is the planar graph for this Platonic solid; such a cycle is designated by the darkened edges. This illustration leads us to the following definition.

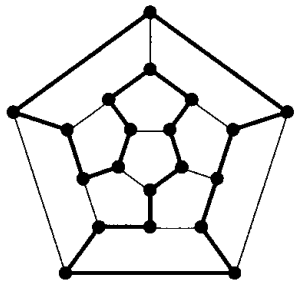


Figure 11.77

Definition 11.21 If $G = (V, E)$ is a graph or multigraph with $|V| \geq 3$, we say that G has a *Hamilton cycle* if there is a cycle in G that contains every vertex in V. A *Hamilton path* is a path (and not a cycle) in G that contains each vertex.

Given a graph with a Hamilton cycle, we find that the deletion of any edge in the cycle results in a Hamilton path. It is possible, however, for a graph to have a Hamilton path without having a Hamilton cycle.

It may seem that the existence of a Hamilton cycle (path) and the existence of an Euler circuit (trail) for a graph are similar problems. The Hamilton cycle (path) is designed to visit each vertex in a graph only once; the Euler circuit (trail) traverses the graph so that each edge is traveled exactly once. Unfortunately, there is no helpful connection between the two ideas, and unlike the situation for Euler circuits (trails), there do not exist necessary

and sufficient conditions on a graph G that guarantee the existence of a Hamilton cycle (path). If a graph has a Hamilton cycle, then it will at least be connected. Many theorems exist that establish either necessary or sufficient conditions for a connected graph to have a Hamilton cycle or path. We shall investigate several of these results later. When confronted with particular graphs, however, we shall often resort to trial and error, with a few helpful observations.

EXAMPLE 11.26

Referring back to the hypercubes in Fig. 11.35 we find in Q_2 the cycle

$$00 \longrightarrow 10 \longrightarrow 11 \longrightarrow 01 \longrightarrow 00$$

and in Q_3 the cycle

$$000 \longrightarrow 100 \longrightarrow 110 \longrightarrow 010 \longrightarrow 011 \longrightarrow 111 \longrightarrow 101 \longrightarrow 001 \longrightarrow 000.$$

Hence Q_2 and Q_3 have Hamilton cycles (and paths). In fact, for all $n \geq 2$, we find that Q_n has a Hamilton cycle. (The reader is asked to establish this in the Section Exercises.) [Note, in addition, that the listings: 00, 10, 11, 01 and 000, 100, 110, 010, 011, 111, 101, 001 are examples of Gray codes (which were introduced in Example 3.9).]

EXAMPLE 11.27

If G is the graph in Fig. 11.78, the edges $\{a, b\}$, $\{b, c\}$, $\{c, f\}$, $\{f, e\}$, $\{e, d\}$, $\{d, g\}$, $\{g, h\}$, $\{h, i\}$ yield a Hamilton path for G. But does G have a Hamilton cycle?

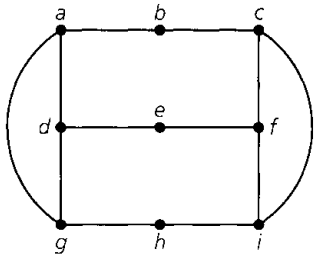


Figure 11.78

Since G has nine vertices, if there is a Hamilton cycle in G it must contain nine edges. Let us start at vertex b and try to build a Hamilton cycle. Because of the symmetry in the graph, it doesn't matter whether we go from b to c or to a. We'll go to c. At c we can go either to f or to i. Using symmetry again, we go to f. Then we delete edge $\{c, i\}$ from further consideration because we cannot return to vertex c. In order to include vertex i in our cycle, we must now go from f to i (to h to g). With edges $\{c, f\}$ and $\{f, i\}$ in the cycle, we cannot have edge $\{e, f\}$ in the cycle. [Otherwise, in the cycle we would have $\deg(f) > 2$.] But then once we get to e we are stuck. Hence there is no Hamilton cycle for the graph.

Example 11.27 indicates a few helpful hints for trying to find a Hamilton cycle in a graph $G = (V, E)$.

1) If G has a Hamilton cycle, then for all $v \in V$, $\deg(v) \geq 2$.
2) If $a \in V$ and $\deg(a) = 2$, then the two edges incident with vertex a must appear in every Hamilton cycle for G.

3) If $a \in V$ and $\deg(a) > 2$, then as we try to build a Hamilton cycle, once we pass through vertex a, any unused edges incident with a are deleted from further consideration.

4) In building a Hamilton cycle for G, we cannot obtain a cycle for a subgraph of G unless it contains all the vertices of G.

Our next example provides an interesting technique for showing that a special type of graph has no Hamilton path.

EXAMPLE 11.28

In Fig. 11.79(a) we have a connected graph G, and we wish to know whether G contains a Hamilton path. Part (b) of the figure provides the same graph with a set of labels x, y. This labeling is accomplished as follows: First we label vertex a with the letter x. Those vertices adjacent to a (namely, b, c, and d) are then labeled with the letter y. Then we label the unlabeled vertices adjacent to b, c, or d with x. This results in the label x on the vertices e, g, and i. Finally, we label the unlabeled vertices adjacent to e, g, or i with the label y. At this point, all the vertices in G are labeled. Now, since $|V| = 10$, if G is to have a Hamilton path there must be an alternating sequence of five x's and five y's. Only four vertices are labeled with x, so this is impossible. Hence G has no Hamilton path (or cycle).

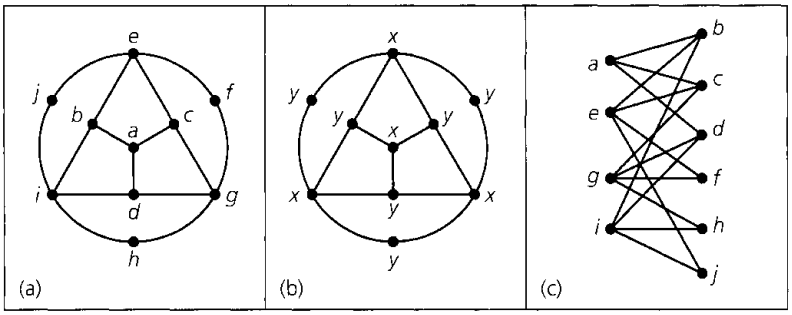


Figure 11.79

But why does this argument work here? In part (c) of Fig. 11.79 we have redrawn the given graph, and we see that it is bipartite. From Exercise 10 in the previous section we know that a bipartite graph cannot have a cycle of odd length. It is also true that if a graph has no cycle of odd length, then it is bipartite. (The proof is requested of the reader in Exercise 9 of this section.) Consequently, whenever a connected graph has no odd cycle (and is bipartite), the method described above may be helpful in determining when the graph does *not* have a Hamilton path. (Exercise 10 in this section examines this idea further.)

Our next example provides an application that calls for Hamilton cycles in a complete graph.

EXAMPLE 11.29

At Professor Alfred's science camp, 17 students have lunch together each day at a circular table. They are trying to get to know one another better, so they make an effort to sit next to two different colleagues each afternoon. For how many afternoons can they do this? How can they arrange themselves on these occasions?

To solve this problem we consider the graph K_n, where $n \geq 3$ and is odd. This graph has n vertices (one for each student) and $\binom{n}{2} = n(n-1)/2$ edges. A Hamilton cycle in K_n

corresponds to a seating arrangement. Each of these cycles has n edges, so we can have at most $(1/n)\binom{n}{2} = (n-1)/2$ Hamilton cycles with no two having an edge in common.

Consider the circle in Fig. 11.80 and the subgraph of K_n consisting of the n vertices and the n edges $\{1, 2\}, \{2, 3\}, \ldots, \{n-1, n\}, \{n, 1\}$. Keep the vertices on the circumference fixed and rotate this Hamilton cycle clockwise through the angle $[1/(n-1)](2\pi)$. This gives us the Hamilton cycle (Fig. 11.81) made up of edges $\{1, 3\}, \{3, 5\}, \{5, 2\}, \{2, 7\}, \ldots, \{n, n-3\}, \{n-3, n-1\}, \{n-1, 1\}$. This Hamilton cycle has no edge in common with the first cycle. When $n \geq 7$ and we continue to rotate the cycle in Fig. 11.80 in this way through angles $[k/(n-1)](2\pi)$, where $2 \leq k \leq (n-3)/2$, we obtain a total of $(n-1)/2$ Hamilton cycles, no two of which have an edge in common.

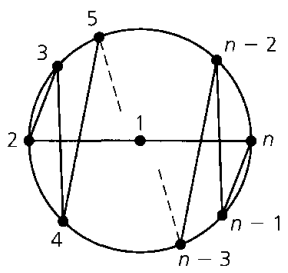


Figure 11.80

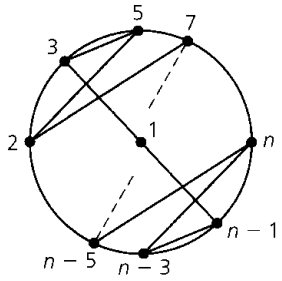


Figure 11.81

Therefore the 17 students at the science camp can dine for $[(17-1)/2] = 8$ days before some student will have to sit next to another student for a second time. Using Fig. 11.80 with $n = 17$, we can obtain eight such possible arrangements.

We turn now to some further results on Hamilton paths and cycles. Our first result was established in 1934 by L. Redei.

THEOREM 11.7

Let K_n^* be a complete directed graph — that is, K_n^* has n vertices and for each distinct pair x, y of vertices, exactly one of the edges (x, y) or (y, x) is in K_n^*. Such a graph (called a *tournament*) always contains a (directed) Hamilton path.

Proof: Let $m \geq 2$ with p_m a path containing the $m-1$ edges $(v_1, v_2), (v_2, v_3), \ldots, (v_{m-1}, v_m)$. If $m = n$, we're finished. If not, let v be a vertex that doesn't appear in p_m.

If (v, v_1) is an edge in K_n^*, we can extend p_m by adjoining this edge. If not, then (v_1, v) must be an edge. Now suppose that (v, v_2) is in the graph. Then we have the larger path: $(v_1, v), (v, v_2), (v_2, v_3), \ldots, (v_{m-1}, v_m)$. If (v, v_2) is not an edge in K_n^*, then (v_2, v) must be. As we continue this process there are only two possibilities: (a) For some $1 \leq k \leq m-1$ the edges $(v_k, v), (v, v_{k+1})$ are in K_n^* and we replace (v_k, v_{k+1}) with this pair of edges; or (b) (v_m, v) is in K_n^* and we add this edge to p_m. Either case results in a path p_{m+1} that includes $m+1$ vertices and has m edges. This process can be repeated until we have such a path p_n on n vertices.

EXAMPLE 11.30

In a round-robin tournament each player plays every other player exactly once. We want to somehow rank the players according to the results of the tournament. Since we could have players a, b, and c where a beats b and b beats c, but c beats a, it is not always possible to have a ranking where a player in a certain position has beaten all of the opponents in

later positions. Representing the players by vertices, construct a directed graph G on these vertices by drawing edge (x, y) if x beats y. Then by Theorem 11.7, it is possible to list the players such that each has beaten the next player on the list.

THEOREM 11.8

Let $G = (V, E)$ be a loop-free graph with $|V| = n \geq 2$. If $\deg(x) + \deg(y) \geq n - 1$ for all $x, y \in V$, $x \neq y$, then G has a Hamilton path.

Proof: First we prove that G is connected. If not, let C_1, C_2 be two components of G and let $x, y \in V$ with x a vertex in C_1 and y a vertex in C_2. Let C_i have n_i vertices, $i = 1, 2$. Then $\deg(x) \leq n_1 - 1$, $\deg(y) \leq n_2 - 1$, and $\deg(x) + \deg(y) \leq (n_1 + n_2) - 2 \leq n - 2$, contradicting the condition given in the theorem. Consequently, G is connected.

Now we build a Hamilton path for G. For $m \geq 2$, let p_m be the path $\{v_1, v_2\}, \{v_2, v_3\}, \ldots, \{v_{m-1}, v_m\}$ of length $m - 1$. (We relabel vertices if necessary.) Such a path exists, because for $m = 2$ all that is needed is one edge. If v_1 is adjacent to any vertex v other than $v_2, v_3, \ldots, v_m$, we add the edge $\{v, v_1\}$ to p_m to get p_{m+1}. The same type of procedure is carried out if v_m is adjacent to a vertex other than $v_1, v_2, \ldots, v_{m-1}$. If we are able to enlarge p_m to p_n in this way, we get a Hamilton path. Otherwise the path p_m: $\{v_1, v_2\}, \ldots, \{v_{m-1}, v_m\}$ has v_1, v_m adjacent only to vertices in p_m, and $m < n$. When this happens we claim that G contains a cycle on these vertices. If v_1 and v_m are adjacent, then the cycle is $\{v_1, v_2\}, \{v_2, v_3\}, \ldots, \{v_{m-1}, v_m\}, \{v_m, v_1\}$. If v_1 and v_m are not adjacent, then v_1 is adjacent to a subset S of the vertices in $\{v_2, v_3, \ldots, v_{m-1}\}$. If there is a vertex $v_t \in S$ such that v_m is adjacent to v_{t-1}, then we can get the cycle by adding $\{v_1, v_t\}, \{v_{t-1}, v_m\}$ to p_m and deleting $\{v_{t-1}, v_t\}$ as shown in Fig. 11.82. If not, let $|S| = k < m - 1$. Then $\deg(v_1) = k$ and $\deg(v_m) \leq (m - 1) - k$, and we have the contradiction $\deg(v_1) + \deg(v_m) \leq m - 1 < n - 1$. Hence there is a cycle connecting $v_1, v_2, \ldots, v_m$.

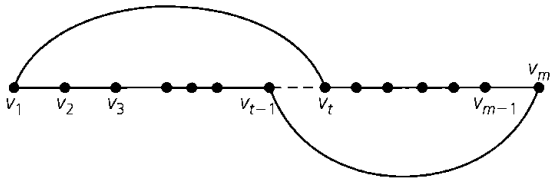


Figure 11.82

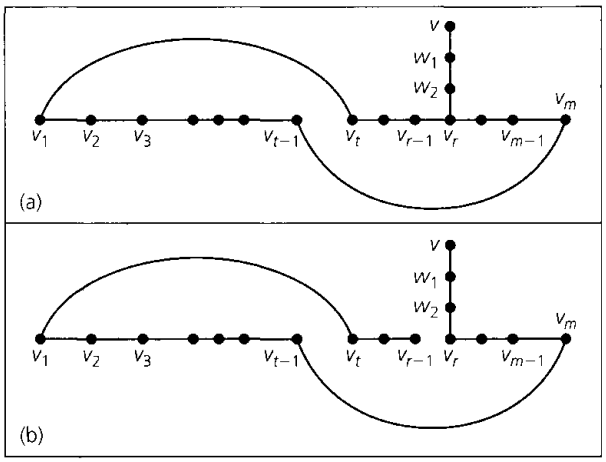


Figure 11.83

Now consider a vertex $v \in V$ that is not found on this cycle. The graph G is connected, so there is a path from v to a first vertex v_r in the cycle, as shown in Fig. 11.83(a). Removing the edge $\{v_{r-1}, v_r\}$ (or $\{v_1, v_t\}$ if $r = t$), we get the path (longer than the original p_m) shown in Fig. 11.83(b). Repeating this process (applied to p_m) for the path in Fig. 11.83(b), we continue to increase the length of the path until it includes every vertex of G.

COROLLARY 11.4 Let $G = (V, E)$ be a loop-free graph with n (≥ 2) vertices. If $\deg(v) \geq (n-1)/2$ for all $v \in V$, then G has a Hamilton path.

Proof: The proof is left as an exercise for the reader.

Our last theorem for this section provides a sufficient condition for the existence of a Hamilton cycle in a loop-free graph. This was first proved by Oystein Ore in 1960.

THEOREM 11.9 Let $G = (V, E)$ be a loop-free undirected graph with $|V| = n \geq 3$. If $\deg(x) + \deg(y) \geq n$ for all nonadjacent $x, y \in V$, then G contains a Hamilton cycle.

Proof: Assume that G does not contain a Hamilton cycle. We add edges to G until we arrive at a subgraph H of K_n, where H has no Hamilton cycle, but, for any edge e (of K_n) not in H, $H + e$ does have a Hamilton cycle.

Since $H \neq K_n$, there are vertices $a, b \in V$, where $\{a, b\}$ is not an edge of H but $H + \{a, b\}$ has a Hamilton cycle C. The graph H has no such cycle, so the edge $\{a, b\}$ is a part of cycle C. Let us list the vertices of H (and G) on cycle C as follows:

$$\overbrace{a\,(=v_1) \to b\,(=v_2) \to v_3 \to v_4 \to \cdots \to v_{n-1} \to v_n}$$

For each $3 \leq i \leq n$, if the edge $\{b, v_i\}$ is in the graph H, then we claim that the edge $\{a, v_{i-1}\}$ cannot be an edge of H. For if both of these edges are in H, for some $3 \leq i \leq n$, then we get the Hamilton cycle

$$\overbrace{b \to v_i \to v_{i+1} \to \cdots \to v_{n-1} \to v_n \to a \to v_{i-1} \to v_{i-2} \to \cdots v_4 \to v_3}$$

for the graph H (which has no Hamilton cycle). Therefore, for each $3 \leq i \leq n$, at most one of the edges $\{b, v_i\}$, $\{a, v_{i-1}\}$ is in H. Consequently,

$$\deg_H(a) + \deg_H(b) < n,$$

where $\deg_H(v)$ denotes the degree of vertex v in graph H. For all $v \in V$, $\deg_H(v) \geq \deg_G(v) = \deg(v)$, so we have nonadjacent (in G) vertices a, b, where

$$\deg(a) + \deg(b) < n.$$

This contradicts the hypothesis that $\deg(x) + \deg(y) \geq n$ for all nonadjacent $x, y \in V$, so we reject our assumption and find that G contains a Hamilton cycle.

Now we shall obtain the following two results from Theorem 11.9. Each will give us a sufficient condition for a loop-free undirected graph $G = (V, E)$ to have a Hamilton cycle. The first result is similar to Corollary 11.4 and is concerned with the degree of each vertex v in V. The second result examines the size of the edge set E.

COROLLARY 11.5 If $G = (V, E)$ is a loop-free undirected graph with $|V| = n \geq 3$, and if $\deg(v) \geq n/2$ for all $v \in V$, then G has a Hamilton cycle.

Proof: We shall leave the proof of this result for the Section Exercises.

COROLLARY 11.6 If $G = (V, E)$ is a loop-free undirected graph with $|V| = n \geq 3$, and if $|E| \geq \binom{n-1}{2} + 2$, then G has a Hamilton cycle.

Proof: Let $a, b \in V$, where $\{a, b\} \notin E$. [Since a, b are nonadjacent, we want to show that $\deg(a) + \deg(b) \geq n$.] Remove the following from the graph G: (i) all edges of the form $\{a, x\}$, where $x \in V$; (ii) all edges of the form $\{y, b\}$, where $y \in V$; and (iii) the vertices a and b. Let $H = (V', E')$ denote the resulting subgraph. Then $|E| = |E'| + \deg(a) + \deg(b)$ because $\{a, b\} \notin E$.

Since $|V'| = n - 2$, H is a subgraph of the complete graph K_{n-2}, so $|E'| \leq \binom{n-2}{2}$. Consequently, $\binom{n-1}{2} + 2 \leq |E| = |E'| + \deg(a) + \deg(b) \leq \binom{n-2}{2} + \deg(a) + \deg(b)$, and we find that

$$\begin{aligned}
\deg(a) + \deg(b) &\geq \binom{n-1}{2} + 2 - \binom{n-2}{2} \\
&= \left(\frac{1}{2}\right)(n-1)(n-2) + 2 - \left(\frac{1}{2}\right)(n-2)(n-3) \\
&= \left(\frac{1}{2}\right)(n-2)[(n-1) - (n-3)] + 2 \\
&= \left(\frac{1}{2}\right)(n-2)(2) + 2 = (n-2) + 2 = n.
\end{aligned}$$

Therefore it follows from Theorem 11.9 that the given graph G has a Hamilton cycle.

A problem that is related to the search for Hamilton cycles in a graph is the *traveling salesman problem*. (An article dealing with this problem was published by Thomas P. Kirkman in 1855.) Here a traveling salesperson leaves his or her home and must visit certain locations before returning. The objective is to find an order in which to visit the locations that is most efficient (perhaps in terms of total distance traveled or total cost). The problem can be modeled with a labeled (edges have distances or costs associated with them) graph where the most efficient Hamilton cycle is sought.

The references by R. Bellman, K. L. Cooke, and J. A. Lockett [7]; M. Bellmore and G. L. Nemhauser [8]; E. A. Elsayed [15]; E. A. Elsayed and R. G. Stern [16]; and L. R. Foulds [17] should prove interesting to the reader who wants to learn more about this important optimization problem. Also, the text edited by E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan, and D. B. Shmoys [22] presents 12 papers on various facets of this problem.

Even more on the traveling salesman problem and its applications can be found in the handbooks edited by M. O. Ball, T. L. Magnanti, C. L. Monma, and G. L. Nemhauser—in particular, the articles by R. K. Ahuja, T. L. Magnanti, J. B. Orlin, and M. R. Reddy [2], and by M. Jünger, G. Reinelt, and G. Rinaldi [21].

EXERCISES 11.5

1. Give an example of a connected graph that has (a) Neither an Euler circuit nor a Hamilton cycle. (b) An Euler circuit but no Hamilton cycle. (c) A Hamilton cycle but no Euler circuit. (d) Both a Hamilton cycle and an Euler circuit.

2. Characterize the type of graph in which an Euler trail (circuit) is also a Hamilton path (cycle).

3. Find a Hamilton cycle, if one exists, for each of the graphs or multigraphs in Fig. 11.84. If the graph has no Hamilton cycle, determine whether it has a Hamilton path.

4. **a)** Show that the Petersen graph [Fig. 11.52(a)] has no Hamilton cycle but that it has a Hamilton path.

b) Show that if any vertex (and the edges incident to it) is removed from the Petersen graph, then the resulting subgraph has a Hamilton cycle.

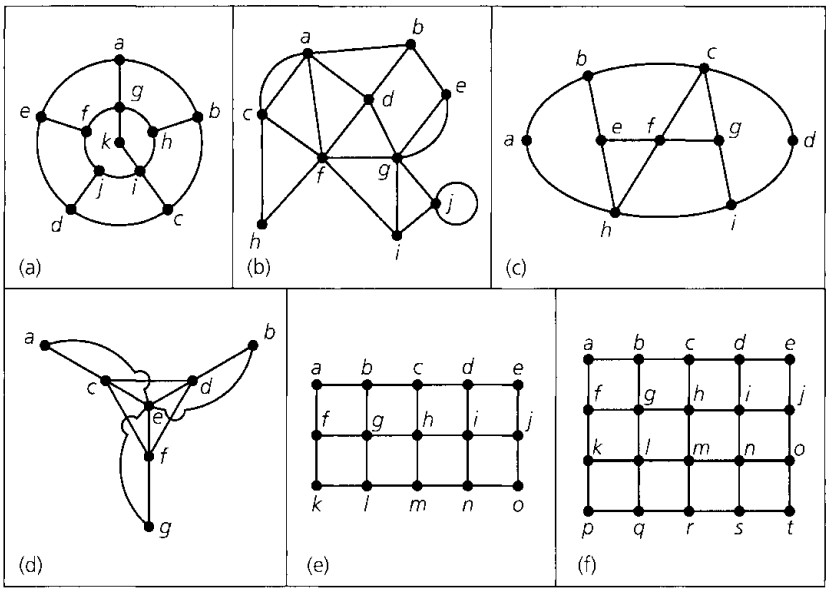


Figure 11.84

5. Consider the graphs in parts (d) and (e) of Fig. 11.84. Is it possible to remove one vertex from each of these graphs so that each of the resulting subgraphs has a Hamilton cycle?

6. If $n \geq 3$, how many different Hamilton cycles are there in the wheel graph W_n? (The graph W_n was defined in Exercise 14 of Section 11.1.)

7. **a)** For $n \geq 3$, how many different Hamilton cycles are there in the complete graph K_n?

b) How many edge-disjoint Hamilton cycles are there in K_{21}?

c) Nineteen students in a nursery school play a game each day where they hold hands to form a circle. For how many days can they do this with no student holding hands with the same playmate twice?

8. **a)** For $n \in \mathbf{Z}^+$, $n \geq 2$, show that the number of distinct Hamilton cycles in the graph $K_{n,n}$ is $(1/2)(n-1)!\,n!$

b) How many different Hamilton paths are there for $K_{n,n}$, $n \geq 1$?

9. Let $G = (V, E)$ be a loop-free undirected graph. Prove that if G contains no cycle of odd length, then G is bipartite.

10. **a)** Let $G = (V, E)$ be a connected bipartite undirected graph with V partitioned as $V_1 \cup V_2$. Prove that if $|V_1| \neq |V_2|$, then G cannot have a Hamilton cycle.

b) Prove that if the graph G in part (a) has a Hamilton path, then $|V_1| - |V_2| = \pm 1$.

c) Give an example of a connected bipartite undirected graph $G = (V, E)$, where V is partitioned as $V_1 \cup V_2$ and $|V_1| = |V_2| - 1$, but G has no Hamilton path.

11. **a)** Determine all nonisomorphic tournaments with three vertices.

b) Find all of the nonisomorphic tournaments with four vertices. List the in degree and the out degree for each vertex, in each of these tournaments.

12. Prove that for $n \geq 2$, the hypercube Q_n has a Hamilton cycle.

13. Let $T = (V, E)$ be a tournament with $v \in V$ of maximum out degree. If $w \in V$ and $w \neq v$, prove that either $(v, w) \in E$ or there is a vertex y in V where $y \neq v, w$, and $(v, y), (y, w) \in E$. (Such a vertex v is called a *king* for the tournament.)

14. Find a counterexample to the converse of Theorem 11.8.

15. Give an example of a loop-free connected undirected multigraph $G = (V, E)$ such that $|V| = n$ and $\deg(x) + \deg(y) \geq n - 1$ for all $x, y \in V$, but G has no Hamilton path.

16. Prove Corollaries 11.4 and 11.5.

17. Give an example to show that the converse of Corollary 11.5 need not be true.

18. Helen and Dominic invite 10 friends to dinner. In this group of 12 people everyone knows at least 6 others. Prove that the 12 can be seated around a circular table in such a way that each person is acquainted with the persons sitting on either side.

19. Let $G = (V, E)$ be a loop-free undirected graph that is 6-regular. Prove that if $|V| = 11$, then G contains a Hamilton cycle.

20. Let $G = (V, E)$ be a loop-free undirected n-regular graph with $|V| \geq 2n + 2$. Prove that $\overline{G}$ (the complement of G) has a Hamilton cycle.

21. For $n \geq 3$, let C_n denote the undirected cycle on n vertices. The graph $\overline{C}_n$, the complement of C_n, is often called the *cocycle* on n vertices. Prove that for $n \geq 5$ the cocycle $\overline{C}_n$ has a Hamilton cycle.

22. Let $n \in \mathbf{Z}^+$ with $n \geq 4$, and let the vertex set V' for the complete graph K_{n-1} be $\{v_1, v_2, v_3, \ldots, v_{n-1}\}$. Now construct the loop-free undirected graph $G_n = (V, E)$ from K_{n-1} as follows: $V = V' \cup \{v\}$, and E consists of all the edges in K_{n-1} except for the edge $\{v_1, v_2\}$, which is replaced by the pair of edges $\{v_1, v\}$ and $\{v, v_2\}$.

a) Determine $\deg(x) + \deg(y)$ for all nonadjacent vertices x and y in V.

b) Does G_n have a Hamilton cycle?

c) How large is the edge set E?

d) Do the results in parts (b) and (c) contradict Corollary 11.6?

23. For $n \in \mathbf{Z}^+$ where $n \geq 4$, let $V' = \{v_1, v_2, v_3, \ldots, v_{n-1}\}$ be the vertex set for the complete graph K_{n-1}. Construct the loop-free undirected graph $H_n = (V, E)$ from K_{n-1} as follows: $V = V' \cup \{v\}$, and E consists of all the edges in K_{n-1} together with the new edge $\{v, v_1\}$.

a) Show that H_n has a Hamilton path but no Hamilton cycle.

b) How large is the edge set E?

24. Let $n = 2^k$ for $k \in \mathbf{Z}^+$. We use the n k-bit sequences (of 0's and 1's) to represent $1, 2, 3, \ldots, n$, so that for two consecutive integers i, $i + 1$, the corresponding k-bit sequences differ in exactly one component. This representation is called a *Gray code* (comparable to what we saw in Example 3.9).

a) For $k = 3$, use a graph model with $V = \{000, 001, 010, \ldots, 111\}$ to find such a code for $1, 2, 3, \ldots, 8$. How is this related to the concept of a Hamilton path?

b) Answer part (a) for $k = 4$.

25. If $G = (V, E)$ is an undirected graph, a subset I of V is called *independent* if no two vertices in I are adjacent. An independent set I is called *maximal* if no vertex v can be added to I with $I \cup \{v\}$ independent. The *independence number* of G, denoted $\beta(G)$, is the size of a largest independent set in G.

a) For each graph in Fig. 11.85 find two maximal independent sets with different sizes.

b) Find $\beta(G)$ for each graph in part (a).

c) Determine $\beta(G)$ for each of the following graphs: (i) $K_{1,3}$; (ii) $K_{2,3}$; (iii) $K_{3,2}$; (iv) $K_{4,4}$; (v) $K_{4,6}$; (vi) $K_{m,n}$, $m, n \in \mathbf{Z}^+$.

d) Let I be an independent set in $G = (V, E)$. What type of subgraph does I induce in $\overline{G}$?

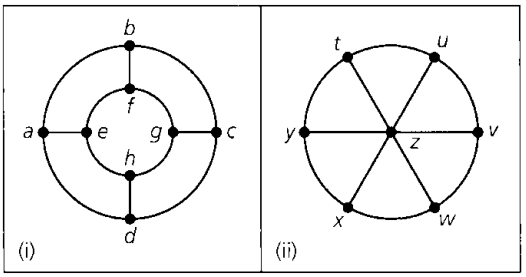


Figure 11.85

26. Let $G = (V, E)$ be an undirected graph with subset I of V an independent set. For each $a \in I$ and each Hamilton cycle C for G, there will be $\deg(a) - 2$ edges in E that are incident with a and not in C. Therefore there are at least $\sum_{a \in I}[\deg(a) - 2] = \sum_{a \in I} \deg(a) - 2|I|$ edges in E that do not appear in C.

a) Why are these $\sum_{a \in I} \deg(a) - 2|I|$ edges distinct?

b) Let $v = |V|$, $e = |E|$. Prove that if

$$e - \sum_{a \in I} \deg(a) + 2|I| < v,$$

then G has no Hamilton cycle.

c) Select a suitable independent set I and use part (b) to show that the graph in Fig. 11.86 (known as the Herschel graph) has no Hamilton cycle.

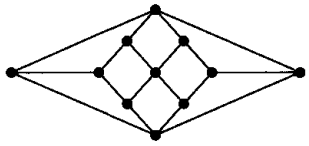


Figure 11.86

11.6 Graph Coloring and Chromatic Polynomials

At the J. & J. Chemical Company, Jeannette is in charge of the storage of chemical compounds in the company warehouse. Since certain types of compounds (such as acids and bases) should not be kept in the same vicinity, she decides to have her partner Jack par-

tition the warehouse into separate storage areas so that incompatible chemical reagents can be stored in separate compartments. How can she determine the number of storage compartments that Jack will have to build?

If this company sells 25 chemical compounds, let $\{c_1, c_2, \ldots, c_{25}\} = V$, a set of vertices. For all $1 \leq i < j \leq 25$, we draw the edge $\{c_i, c_j\}$ if c_i and c_j must be stored in separate compartments. This gives us an undirected graph $G = (V, E)$.

We now introduce the following concept.

Definition 11.22 If $G = (V, E)$ is an undirected graph, a *proper coloring* of G occurs when we color the vertices of G so that if $\{a, b\}$ is an edge in G, then a and b are colored with different colors. (Hence adjacent vertices have different colors.) The minimum number of colors needed to properly color G is called the *chromatic number* of G and is written $\chi(G)$.

Returning to assist Jeannette at the warehouse, we find that the number of storage compartments Jack must build is equal to $\chi(G)$ for the graph we constructed on $V = \{c_1, c_2, \ldots, c_{25}\}$. But how do we compute $\chi(G)$? Before we present any work on how to determine the chromatic number of a graph, we turn to the following related idea.

In Example 11.24 we mentioned the connection between coloring the regions in a planar map (with neighboring regions having different colors) and properly coloring the vertices in an associated graph. Determining the smallest number of colors needed to color planar maps in this way has been a problem of interest for over a century.

In about 1850, Francis Guthrie (1831–1899) became interested in the general problem after showing how to color the counties on a map of England with only four colors. Shortly thereafter, he showed the "Four-color Problem" to his younger brother Frederick (1833–1866), who was then a student of Augustus DeMorgan (1806–1871). DeMorgan communicated the problem (in 1852) to William Hamilton (1805–1865). The problem did not interest Hamilton and lay dormant for about 25 years. Then, in 1878, the scientific community was made aware of the problem through an announcement by Arthur Cayley (1821–1895) at a meeting of the London Mathematical Society. In 1879 Cayley stated the problem in the first volume of the *Proceedings of the Royal Geographical Society*. Shortly thereafter, the British barrister (and keen amateur mathematician) Sir Alfred Kempe (1849–1922) devised a proof that remained unquestioned for over a decade. In 1890, however, the British mathematician Percy John Heawood (1861–1955) found a mistake in Kempe's work.

The problem remained unsolved until 1976, when it was finally settled by Kenneth Appel and Wolfgang Haken. Their proof employs a very intricate computer analysis of 1936 (reducible) configurations.

Although only four colors are needed to properly color the regions in a planar map, we need more than four colors to properly color the vertices of some nonplanar graphs.

We start with some small examples. Then we shall find a way to determine $\chi(G)$ from smaller subgraphs of G — in certain situations. [In general, computing $\chi(G)$ is a very difficult problem.] We shall also obtain what is called the chromatic polynomial for G and see how it can be used in computing $\chi(G)$.

EXAMPLE 11.31 For the graph G in Fig. 11.87, we start at vertex a and next to each vertex write the number of a color needed to properly color the vertices of G that have been considered up to that point. Going to vertex b, the 2 indicates the need for a second color because vertices a and b are adjacent. Proceeding alphabetically to f, we find that two colors are needed to

properly color $\{a, b, c, d, e, f\}$. For vertex g a third color is needed; this third color can also be used for vertex h because $\{g, h\}$ is not an edge in G. Thus this sequential coloring (labeling) method gives us a proper coloring for G, so $\chi(G) \leq 3$. Since K_3 is a subgraph of G [for example, the subgraph induced by a, b and g is (isomorphic to) K_3], we have $\chi(G) \geq 3$, so $\chi(G) = 3$.

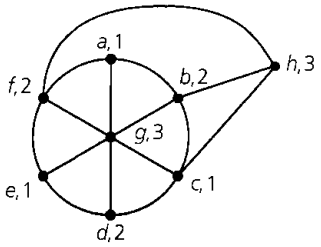


Figure 11.87

EXAMPLE 11.32

a) For all $n \geq 1$, $\chi(K_n) = n$.

b) The chromatic number of the Herschel graph (Fig. 11.86) is 2.

c) If G is the Petersen graph [see Fig. 11.52 (a)], then $\chi(G) = 3$.

EXAMPLE 11.33

Let G be the graph shown in Fig. 11.88. For $U = \{b, f, h, i\}$, the induced subgraph $\langle U \rangle$ of G is isomorphic to K_4, so $\chi(G) \geq \chi(K_4) = 4$. Therefore, if we can determine a way to properly color the vertices of G with four colors, then we shall know that $\chi(G) = 4$. One way to accomplish this is to color the vertices e, f, g blue; the vertices b, j red; the vertices c, h white; and the vertices a, d, i green.

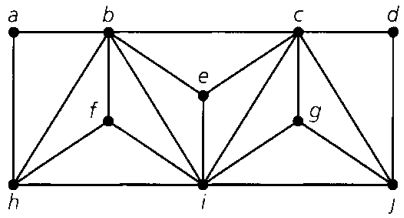


Figure 11.88

We turn now to a method for determining $\chi(G)$. Our coverage follows the development in the survey article [25] by R. C. Read.

Let G be an undirected graph, and let λ be the number of colors that we have available for properly coloring the vertices of G. Our objective is to find a polynomial function $P(G, \lambda)$, in the variable λ, called the *chromatic polynomial* of G, that will tell us in how many different ways we can properly color the vertices of G, using at most λ colors.

Throughout this discussion, the vertices in an undirected graph $G = (V, E)$ are distinguished by labels. Consequently, two proper colorings of such a graph will be considered different in the following sense: A proper coloring (of the vertices of G) that uses at most λ colors is a function f, with domain V and codomain $\{1, 2, 3, \ldots, \lambda\}$, where $f(u) \neq f(v)$,

for adjacent vertices $u, v \in V$. Proper colorings are then different in the same way that these functions are different.

EXAMPLE 11.34

a) If $G = (V, E)$ with $|V| = n$ and $E = \emptyset$, then G consists of n isolated points, and by the rule of product, $P(G, \lambda) = \lambda^n$.

b) If $G = K_n$, then at least n colors must be available for us to color G properly. Here, by the rule of product, $P(G, \lambda) = \lambda(\lambda - 1)(\lambda - 2) \cdots (\lambda - n + 1)$, which we denote by $\lambda^{(n)}$. For $\lambda < n$, $P(G, \lambda) = 0$ and there are no ways to properly color K_n. $P(G, \lambda) > 0$ for the first time when $\lambda = n = \chi(G)$.

c) For each path in Fig. 11.89, we consider the number of choices (of the λ colors) at each successive vertex. Proceeding alphabetically, we find that $P(G_1, \lambda) = \lambda(\lambda - 1)^3$ and $P(G_2, \lambda) = \lambda(\lambda - 1)^4$. Since $P(G_1, 1) = 0 = P(G_2, 1)$, but $P(G_1, 2) = 2 = P(G_2, 2)$, it follows that $\chi(G_1) = \chi(G_2) = 2$. If five colors are available we can properly color G_1 in $5(4)^3 = 320$ ways; G_2 can be so colored in $5(4)^4 = 1280$ ways.

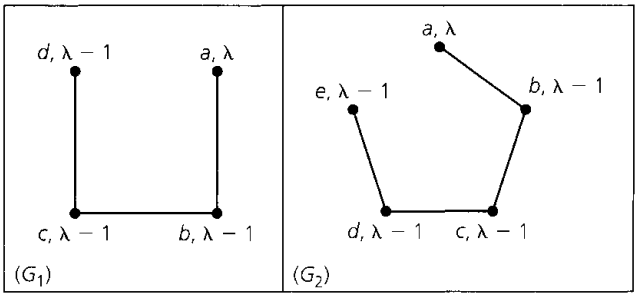


Figure 11.89

In general, if G is a path on n vertices, then $P(G, \lambda) = \lambda(\lambda - 1)^{n-1}$.

d) If G is made up of components $G_1, G_2, \ldots, G_k$, then again by the rule of product, it follows that $P(G, \lambda) = P(G_1, \lambda) \cdot P(G_2, \lambda) \cdots P(G_k, \lambda)$.

As a result of Example 11.34(d), we shall concentrate on connected graphs. In many instances in discrete mathematics, methods have been employed to solve problems in large cases by breaking these down into two or more smaller cases. Once again we use this method of attack. To do so, we need the following ideas and notation.

Let $G = (V, E)$ be an undirected graph. For $e = \{a, b\} \in E$, let G_e denote the subgraph of G obtained by deleting e from G, without removing vertices a and b; that is, $G_e = G - e$ as defined in Section 11.2. From G_e a second subgraph of G is obtained by coalescing (or, identifying) the vertices a and b. This second subgraph is denoted by G'_e.

EXAMPLE 11.35

Figure 11.90 shows G_e and G'_e for graph G with the edge e as specified. Note how the coalescing of a and b in G'_e results in the coalescing of the two pairs of edges $\{d, b\}$, $\{d, a\}$ and $\{a, c\}$, $\{b, c\}$.

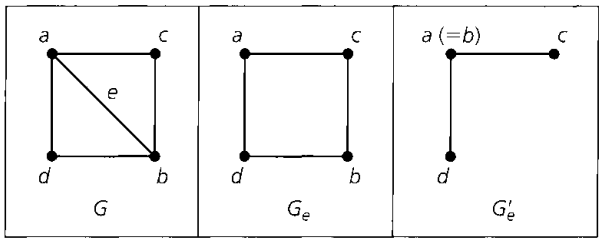


Figure 11.90

Using these special subgraphs, we turn now to the main result.

THEOREM 11.10

Decomposition Theorem for Chromatic Polynomials. If $G = (V, E)$ is a connected graph and $e \in E$, then

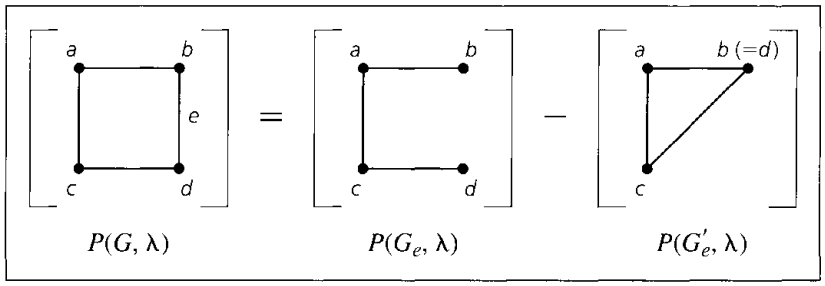
$$P(G_e, \lambda) = P(G, \lambda) + P(G'_e, \lambda).$$

Proof: Let $e = \{a, b\}$. The number of ways to properly color the vertices in G_e with (at most) λ colors is $P(G_e, \lambda)$. Those colorings where a and b have different colors are proper colorings of G. The colorings of G_e that are not proper colorings of G occur when a and b have the same color. But each of these colorings corresponds with a proper coloring for G'_e. This partition of the $P(G_e, \lambda)$ proper colorings of G_e into two disjoint subsets results in the formula $P(G_e, \lambda) = P(G, \lambda) + P(G'_e, \lambda)$.

When calculating chromatic polynomials, we shall place brackets about a graph to indicate its chromatic polynomial.

EXAMPLE 11.36

The following calculations yield $P(G, \lambda)$ for G a cycle of length 4.



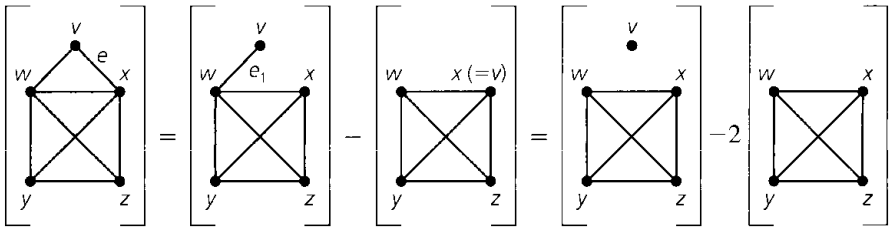
From Example 11.34(c) it follows that $P(G_e, \lambda) = \lambda(\lambda - 1)^3$. With $G'_e = K_3$ we have $P(G'_e, \lambda) = \lambda^{(3)}$. Therefore,

$$P(G, \lambda) = \lambda(\lambda - 1)^3 - \lambda(\lambda - 1)(\lambda - 2) = \lambda(\lambda - 1)[(\lambda - 1)^2 - (\lambda - 2)]$$
$$= \lambda(\lambda - 1)[\lambda^2 - 3\lambda + 3] = \lambda^4 - 4\lambda^3 + 6\lambda^2 - 3\lambda.$$

Since $P(G, 1) = 0$ while $P(G, 2) = 2 > 0$, we know that $\chi(G) = 2$.

EXAMPLE 11.37

Here we find a second application of Theorem 11.10.



$$= \underbrace{(\lambda)(\lambda^{(4)})} - 2\lambda^{(4)} = (\lambda - 2)\lambda^{(4)} = \lambda(\lambda - 1)(\lambda - 2)^2(\lambda - 3)$$

For the disconnected graph with the components K_1, K_4

For each $1 \leq \lambda \leq 3$, $P(G, \lambda) = 0$, but $P(G, \lambda) > 0$ for all $\lambda \geq 4$. Consequently, the given graph has chromatic number 4.

The chromatic polynomials given in Examples 11.36 and 11.37 suggest the following results.

THEOREM 11.11

For each graph G, the constant term in $P(G, \lambda)$ is 0.

Proof: For each graph G, $\chi(G) > 0$ because $V \neq \emptyset$. If $P(G, \lambda)$ has constant term a, then $P(G, 0) = a \neq 0$. This implies that there are a ways to color G properly with 0 colors, a contradiction.

THEOREM 11.12

Let $G = (V, E)$ with $|E| > 0$. Then the sum of the coefficients in $P(G, \lambda)$ is 0.

Proof: Since $|E| \geq 1$, we have $\chi(G) \geq 2$, so we cannot properly color G with only one color. Consequently, $P(G, 1) = 0 =$ the sum of the coefficients in $P(G, \lambda)$.

Since the chromatic polynomial of a complete graph is easy to determine, an alternative method for finding $P(G, \lambda)$ can be obtained. Theorem 11.10 reduced the problem to smaller graphs. Here we add edges to a given graph until we reach complete graphs.

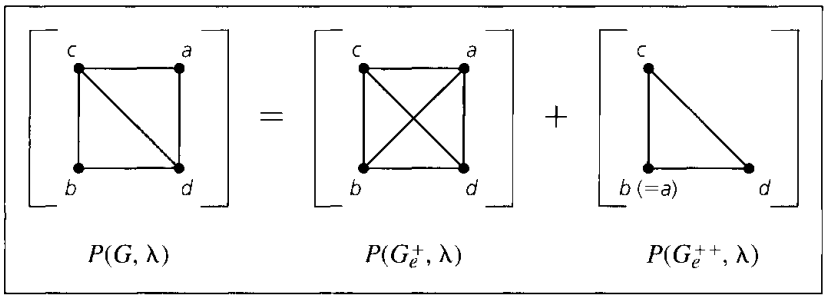
THEOREM 11.13

Let $G = (V, E)$, with $a, b \in V$ but $\{a, b\} = e \notin E$. We write G_e^+ for the graph we obtain from G by adding the edge $e = \{a, b\}$. Coalescing the vertices a and b in G gives us the subgraph G_e^{++} of G. Under these circumstances $P(G, \lambda) = P(G_e^+, \lambda) + P(G_e^{++}, \lambda)$.

Proof: This result follows as in Theorem 11.10 because $P(G_e^+, \lambda) = P(G, \lambda) - P(G_e^{++}, \lambda)$.

EXAMPLE 11.38

Let us now apply Theorem 11.13.



Here $P(G, \lambda) = \lambda^{(4)} + \lambda^{(3)} = \lambda(\lambda - 1)(\lambda - 2)^2$, so $\chi(G) = 3$. In addition, if six colors are available, the vertices in G can be properly colored in $6(5)(4)^2 = 480$ ways.

Our next result again uses complete graphs — along with the following concepts. For all graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$.

i) the *union* of G_1 and G_2, denoted $G_1 \cup G_2$, is the graph with vertex set $V_1 \cup V_2$ and edge set $E_1 \cup E_2$; and

ii) when $V_1 \cap V_2 \neq \emptyset$, the *intersection* of G_1 and G_2, denoted $G_1 \cap G_2$, is the graph with vertex set $V_1 \cap V_2$ and edge set $E_1 \cap E_2$.

THEOREM 11.14

Let G be an undirected graph with subgraphs G_1, G_2. If $G = G_1 \cup G_2$ and $G_1 \cap G_2 = K_n$, for some $n \in \mathbf{Z}^+$, then

$$P(G, \lambda) = \frac{P(G_1, \lambda) \cdot P(G_2, \lambda)}{\lambda^{(n)}}.$$

Proof: Since $G_1 \cap G_2 = K_n$, it follows that K_n is a subgraph of both G_1 and G_2 and that $\chi(G_1), \chi(G_2) \geq n$. Given λ colors, there are $\lambda^{(n)}$ proper colorings of K_n. For each of these $\lambda^{(n)}$ colorings there are $P(G_1, \lambda)/\lambda^{(n)}$ ways to properly color the remaining vertices in G_1. Likewise, there are $P(G_2, \lambda)/\lambda^{(n)}$ ways to properly color the remaining vertices in G_2. By the rule of product,

$$P(G, \lambda) = P(K_n, \lambda) \cdot \frac{P(G_1, \lambda)}{\lambda^{(n)}} \cdot \frac{P(G_2, \lambda)}{\lambda^{(n)}} = \frac{P(G_1, \lambda) \cdot P(G_2, \lambda)}{\lambda^{(n)}}.$$

EXAMPLE 11.39

Consider the graph in Example 11.37. Let G_1 be the subgraph induced by the vertices w, x, y, z. Let G_2 be the complete graph K_3 — with vertices v, w, and x. Then $G_1 \cap G_2$ is the edge $\{w, x\}$, so $G_1 \cap G_2 = K_2$.

Therefore

$$\begin{aligned} P(G, \lambda) &= \frac{P(G_1, \lambda) \cdot P(G_2, \lambda)}{\lambda^{(2)}} = \frac{\lambda^{(4)} \cdot \lambda^{(3)}}{\lambda^{(2)}} \\ &= \frac{\lambda^2(\lambda - 1)^2(\lambda - 2)^2(\lambda - 3)}{\lambda(\lambda - 1)} \\ &= \lambda(\lambda - 1)(\lambda - 2)^2(\lambda - 3), \end{aligned}$$

agreeing with the answer obtained in Example 11.37.

Much more can be said about chromatic polynomials — in particular, there are many unanswered questions. For example, no one has found a set of conditions that indicate whether a given polynomial in λ is the chromatic polynomial for some graph. More about this topic is introduced in the article by R. C. Read [25].

EXERCISES 11.6

1. A pet-shop owner receives a shipment of tropical fish. Among the different species in the shipment are certain pairs where one species feeds on the other. These pairs must consequently be kept in different aquaria. Model this problem as a graph-coloring problem, and tell how to determine the smallest number of aquaria needed to preserve all the fish in the shipment.

2. As the chair for church committees, Mrs. Blasi is faced with scheduling the meeting times for 15 committees. Each committee meets for one hour each week. Two committees having a common member must be scheduled at different times. Model this problem as a graph-coloring problem, and tell how to determine the least number of meeting times Mrs. Blasi has to consider for scheduling the 15 committee meetings.

3. **a)** At the J. & J. Chemical Company, Jeannette has received three shipments that contain a total of seven different chemicals. Furthermore, the nature of these chemicals is such that for all $1 \leq i \leq 5$, chemical i cannot be stored in the same storage compartment as chemical $i + 1$ or chemical $i + 2$. Determine the smallest number of separate storage compartments that Jeannette will need to safely store these seven chemicals.

b) Suppose that in addition to the conditions in part (a), the following four pairs of these same seven chemicals also require separate storage compartments: 1 and 4, 2 and 5, 2 and 6, and 3 and 6. What is the smallest number of storage compartments that Jeannette now needs to safely store the seven chemicals?

4. Give an example of an undirected graph $G = (V, E)$, where $\chi(G) = 3$ but no subgraph of G is isomorphic to K_3.

5. **a)** Determine $P(G, \lambda)$ for $G = K_{1,3}$.

b) For $n \in \mathbf{Z}^+$, what is the chromatic polynomial for $K_{1,n}$? What is its chromatic number?

6. **a)** Consider the graph $K_{2,3}$ shown in Fig. 11.91, and let $\lambda \in \mathbf{Z}^+$ denote the number of colors available to properly color the vertices of $K_{2,3}$. (i) How many proper colorings of $K_{2,3}$ have vertices a, b colored the same? (ii) How many proper colorings of $K_{2,3}$ have vertices a, b colored with different colors?

b) What is the chromatic polynomial for $K_{2,3}$? What is $\chi(K_{2,3})$?

c) For $n \in \mathbf{Z}^+$, what is the chromatic polynomial for $K_{2,n}$? What is $\chi(K_{2,n})$?

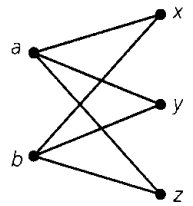


Figure 11.91

7. Find the chromatic number of the following graphs.

a) The complete bipartite graphs $K_{m,n}$.

b) A cycle on n vertices, $n \geq 3$.

c) The graphs in Figs. 11.59(d), 11.62(a), and 11.85.

d) The n-cube Q_n, $n \geq 1$.

8. If G is a loop-free undirected graph with at least one edge, prove that G is bipartite if an only if $\chi(G) = 2$.

9. **a)** Determine the chromatic polynomials for the graphs in Fig. 11.92

b) Find $\chi(G)$ for each graph.

c) If five colors are available, in how many ways can the vertices of each graph be properly colored?

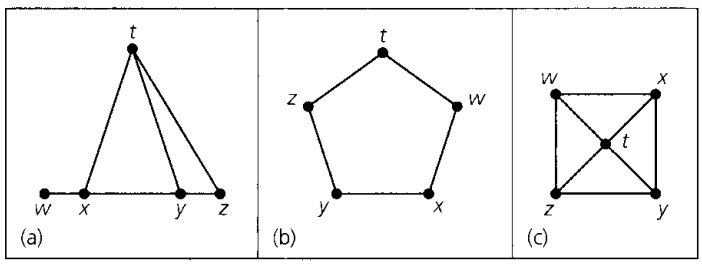


Figure 11.92

10. **a)** Determine whether the graphs in Fig. 11.93 are isomorphic.

b) Find $P(G, \lambda)$ for each graph.

c) Comment on the results found in parts (a) and (b).

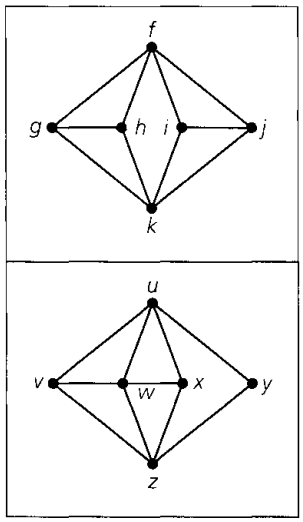


Figure 11.93

11. For $n \geq 3$, let $G_n = (V, E)$ be the undirected graph obtained from the complete graph K_n upon deletion of one edge. Determine $P(G_n, \lambda)$ and $\chi(G_n)$.

12. Consider the complete graph K_n for $n \geq 3$. Color r of the vertices in K_n red and the remaining $n - r\ (= g)$ vertices green. For any two vertices v, w in K_n color the edge $\{v, w\}$ (1) red if v, w are both red; (2) green if v, w are both green; or (3) blue if v, w have different colors. Assume that $r \geq g$.

a) Show that for $r = 6$ and $g = 3$ (and $n = 9$) the total number of red and green edges in K_9 equals the number of blue edges in K_9.

b) Show that the total number of red and green edges in K_n equals the number of blue edges in K_n if and only if $n = r + g$, where g, r are consecutive triangular numbers. [The triangular numbers are defined recursively by $t_1 = 1$, $t_{n+1} = t_n + (n + 1)$, $n \geq 1$; so $t_n = n(n + 1)/2$. Hence $t_1 = 1, t_2 = 3, t_3 = 6, \ldots$.]

13. Let $G = (V, E)$ be the undirected connected "ladder graph" shown in Fig. 11.94.

a) Determine $|V|$ and $|E|$.

b) Prove that $P(G, \lambda) = \lambda(\lambda - 1)(\lambda^2 - 3\lambda + 3)^{n-1}$.

14. Let G be a loop-free undirected graph, where $\Delta = \max_{v \in V}\{\deg(v)\}$. (a) Prove that $\chi(G) \leq \Delta + 1$. (b) Find two types of graphs G, where $\chi(G) = \Delta + 1$.

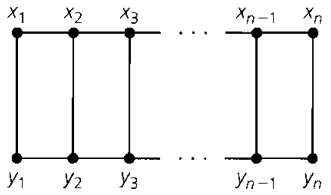


Figure 11.94

15. For $n \geq 3$, let C_n denote the cycle of length n.

a) What is $P(C_3, \lambda)$?

b) If $n \geq 4$, show that

$$P(C_n, \lambda) = P(P_{n-1}, \lambda) - P(C_{n-1}, \lambda),$$

where P_{n-1} denotes the path of length $n - 1$.

c) Verify that $P(P_{n-1}, \lambda) = \lambda(\lambda - 1)^{n-1}$, for all $n \geq 2$.

d) Establish the relations

$$P(C_n, \lambda) - (\lambda - 1)^n = (\lambda - 1)^{n-1} - P(C_{n-1}, \lambda), \qquad n \geq 4,$$

$$P(C_n, \lambda) - (\lambda - 1)^n = P(C_{n-2}, \lambda) - (\lambda - 1)^{n-2}, \qquad n \geq 5.$$

e) Prove that for all $n \geq 3$,

$$P(C_n, \lambda) = (\lambda - 1)^n + (-1)^n(\lambda - 1).$$

16. For $n \geq 3$, recall that the wheel graph, W_n, is obtained from a cycle of length n by placing a new vertex within the cycle and adding edges (spokes) from this new vertex to each vertex of the cycle.

a) What relationship is there between $\chi(C_n)$ and $\chi(W_n)$?

b) Use part (e) of Exercise 15 to show that

$$P(W_n, \lambda) = \lambda(\lambda - 2)^n + (-1)^n\lambda(\lambda - 2).$$

c) **i)** If we have k different colors available, in how many ways can we paint the walls and ceiling of a pentagonal room if adjacent walls, and any wall and the ceiling, are to be painted with different colors?

ii) What is the smallest value of k for which such a coloring is possible?

17. Let $G = (V, E)$ be a loop-free undirected graph with chromatic polynomial $P(G, \lambda)$ and $|V| = n$. Use Theorem 11.13 to prove that $P(G, \lambda)$ has degree n and leading coefficient 1 (that is, the coefficient of λ^n is 1).

18. Let $G = (V, E)$ be a loop-free undirected graph.

a) For each such graph, where $|V| \leq 3$, find $P(G, \lambda)$ and show that in it the terms contain consecutive powers of λ. Also show that the coefficients of these consecutive powers alternate in sign.

b) Now consider $G = (V, E)$, where $|V| = n \geq 4$ and $|E| = k$. Prove by mathematical induction that the terms in $P(G, \lambda)$ contain consecutive powers of λ and that the coefficients of these consecutive powers alternate in sign. [For the induction hypothesis, assume that the result is

true for all loop-free undirected graphs $G = (V, E)$, where either (i) $|V| = n - 1$ or (ii) $|V| = n$, but $|E| = k - 1$.]

c) Prove that if $|V| = n$, then the coefficient of λ^{n-1} in $P(G, \lambda)$ is the negative of $|E|$.

19. Let $G = (V, E)$ be a loop-free undirected graph. We call G *color-critical* if $\chi(G) > \chi(G - v)$ for all $v \in V$.

a) Explain why cycles with an odd number of vertices are color-critical while cycles with an even number of vertices are not color-critical.

b) For $n \in \mathbf{Z}^+$, $n \geq 2$, which of the complete graph K_n are color-critical?

c) Prove that a color-critical graph must be connected.

d) Prove that if G is color-critical with $\chi(G) = k$, then $\deg(v) \geq k - 1$ for all $v \in V$.

11.7 Summary and Historical Review

Unlike other areas in mathematics, graph theory traces its beginnings to a definite time and place: the problem of the seven bridges of Königsberg, which was solved in 1736 by Leonhard Euler (1707–1783). And in 1752 we find Euler's Theorem for planar graphs. (This result was originally presented in terms of polyhedra.) However, after these developments, little was accomplished in this area for almost a century.

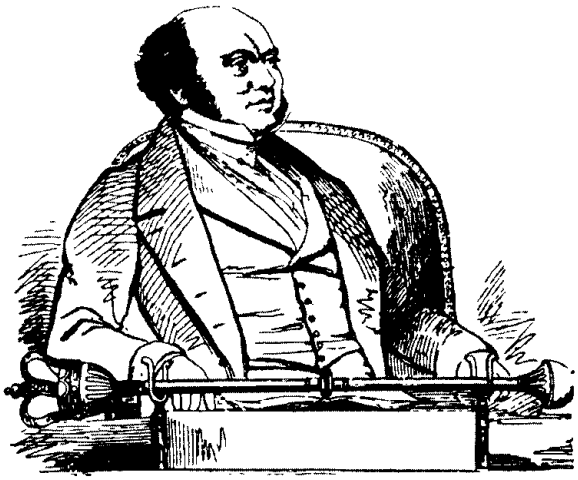
Then, in 1847, Gustav Kirchhoff (1824–1887) examined a special type of graph called a tree. (A *tree* is a loop-free undirected graph that is connected but contains no cycles.) Kirchhoff used this concept in applications dealing with electrical networks in his extension of Ohm's laws for electrical flow. Ten years later Arthur Cayley (1821–1895) developed this same type of graph in order to count the distinct isomers of the saturated hydrocarbons $C_n H_{2n+2}$, $n \in \mathbf{Z}^+$.

This period also saw two other major ideas come to light. The *four-color conjecture* was first investigated by Francis Guthrie (1831–1899) in about 1850. In Section 11.6 we related some of the history of this problem, which was solved via an intricate computer analysis in 1976 by Kenneth Appel and Wolfgang Haken.

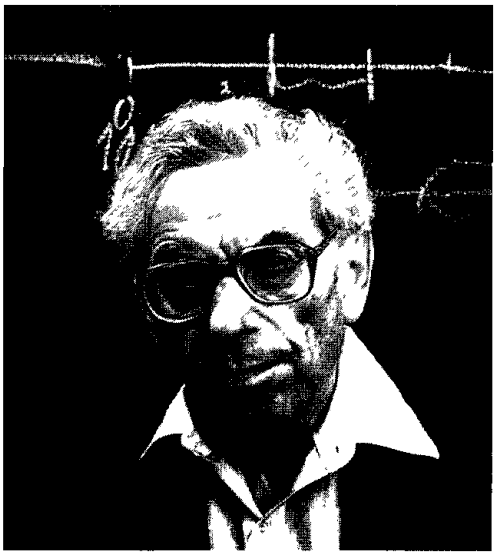
The second major idea was the Hamilton cycle. This cycle is named for Sir William Rowan Hamilton (1805–1865), who used the idea in 1859 for an intriguing puzzle that used the edges on a regular dodecahedron. A solution to this puzzle is not very difficult to find, but mathematicians still search for necessary and sufficient conditions to characterize those undirected graphs that possess a Hamilton path or cycle.

Following these developments, we find little activity until after 1920. The characterization of planar graphs was solved by the Polish mathematician Kasimir Kuratowski (1896–1980) in 1930. In 1936 we find the publication of the first book on graph theory, written by the Hungarian mathematician Dénes König (1884–1944), a prominent researcher in the field. Since then there has been a great deal of activity in the area, the computer providing assistance in the last five decades. Among the many contemporary researchers (not mentioned in the chapter references) in this and related fields one finds the names Claude Berge, V. Chvátal, Paul Erdös, Laszlo Lovász, W. T. Tutte, and Hassler Whitney.

Comparable coverage of the material presented in this chapter is contained in Chapters 6, 8, and 9 of C. L. Liu [23]. More advanced work is found in the works by J. A. Bondy and U. S. R. Murty [10], N. Hartsfield and G. Ringel [20], and D. B. West [32]. The book by F. Buckley and F. Harary [11] revises the classic work of F. Harary [18] and brings the reader up to date on the topics covered in the original 1969 work. The text by G. Chartrand and L. Lesniak [12] provides a more algorithmic approach in its presentation. A proof of



William Rowan Hamilton (1805–1865)
Reproduced courtesy of The Granger Collection, New York



Paul Erdös (1913–1996)
Reproduced courtesy of Christopher Barker

Kuratowski's Theorem appears in Chapter 8 of C. L. Liu [23] and Chapter 6 of D. B. West [32]. The article by G. Chartrand and R. J. Wilson [13] develops many important concepts in graph theory by focusing on one particular graph — the Petersen graph. This graph (which we mentioned in Section 11.4) is named for the Danish mathematician Julius Peter Christian Petersen (1839–1910), who discussed the graph in a paper in 1898.

Applications of graph theory in electrical networks can be found in S. Seshu and M. B. Reed [30]. In the text by N. Deo [14], applications in coding theory, electrical networks, operations research, computer programming, and chemistry occupy Chapters 12–15. The text by F. S. Roberts [26] applies the methods of graph theory to the social sciences. Applications of graph theory in chemistry are given in the article by D. H. Rouvray [29].

More on chromatic polynomials can be found in the survey article by R. C. Read [25]. The role of Polya's theory[†] in graphical enumeration is examined in Chapter 10 of N. Deo [14]. A thorough coverage of this topic is found in the text by F. Harary and E. M. Palmer [19].

Additional coverage on the historical development of graph theory is given in N. Biggs, E. K. Lloyd, and R. J. Wilson [9].

Many applications in graph theory involve large graphs that require the computationally intensive talents of a computer in conjunction with the ingenuity of mathematical methods. Chapter 11 of N. Deo [14] presents computer algorithms dealing with several of the graph-theoretic properties we have studied here. Along the same line, the text by A. V. Aho, J. E. Hopcroft, and J. D. Ullman [1] provides even more for the reader interested in computer science.

As mentioned at the end of Section 11.5, the traveling salesman problem is closely related to the search for a Hamilton cycle in a graph. This is a graph-theoretic problem of interest in both operations research and computer science. The article by M. Bellmore and G. L.

[†]We shall introduce the basic ideas behind this method of enumeration in Chapter 16.

Nemhauser [8] provides a good introductory survey of results on this problem. The text by R. Bellman, K. L. Cooke, and J. A. Lockett [7] includes an algorithmic treatment of this problem along with other graph problems. A number of heuristics for obtaining an approximate solution to the problem are given in Chapter 4 of the text by L. R. Foulds [17]. The text edited by E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan, and D. B. Shmoys [22] contains 12 papers dealing with various aspects of this problem, including historical considerations as well as some results on computational complexity. Applications, where a robot visits different locations in an automated warehouse in order to fill a given order, are examined in the articles by E. A. Elsayed [15] and by E. A. Elsayed and R. G. Stern [16].

The solution of the four-color problem can be examined further by starting with the paper by K. Appel and W. Haken [3]. The problem, together with its history and solution, is examined in the text by D. Barnette [6] and in the *Scientific American* article by K. Appel and W. Haken [4]. The proof uses a computer analysis to handle a large number of cases; the article by T. Tymoczko [31] examines the role of such techniques in pure mathematics. In [5] K. Appel and W. Haken further examine their proof in the light of the computer analysis that was used. The articles by N. Robertson, D. P. Sanders, P. D. Seymour, and R. Thomas [27, 28] provide a simplified proof. In 1997 their computer code was made available on the Internet. This code could prove the four-color problem on a desktop workstation in roughly three hours.

Finally, the article by A. Ralston [24] demonstrates some of the connections among coding theory, combinatorics, graph theory, and computer science.

REFERENCES

1. Aho, Alfred V., Hopcroft, John E., and Ullman, Jeffrey D. *Data Structures and Algorithms.* Reading, Mass.: Addison-Wesley, 1983.
2. Ahuja, Ravindra K., Magnanti, Thomas L., Orlin, James B., and Reddy, M. R."Applications of Network Optimization." In M. O. Ball, Thomas L. Magnanti, C. L. Monma, and G. L. Nemhauser, eds., *Handbooks in Operations Research and Management Science,* Vol. 7, *Network Models.* Amsterdam, Holland: Elsevier, 1995, pp. 1–83.
3. Appel, Kenneth, and Haken, Wolfgang. "Every Planar Map Is Four Colorable." *Bulletin of the American Mathematical Society* 82 (1976): pp. 711–712.
4. Appel, Kenneth, and Haken, Wolfgang."The Solution of the Four-Color-Map Problem. " *Scientific American* 237 (October 1977): pp. 108–121.
5. Appel, Kenneth, and Haken, Wolfgang."The Four Color Proof Suffices." *Mathematical Intelligencer* 8, no. 1 (1986): pp. 10– 20.
6. Barnette, David. *Map Coloring, Polyhedra, and the Four-Color Problem.* Washington, D.C.: The Mathematical Association of America, 1983.
7. Bellman, R., Cooke, K. L., and Lockett, J. A. *Algorithms, Graphs, and Computers.* New York: Academic Press, 1970.
8. Bellmore, M., and Nemhauser, G. L."The Traveling Salesman Problem: A Survey." *Operations Research* 16 (1968): pp. 538–558.
9. Biggs, N., Lloyd, E. K., and Wilson, R. J. *Graph Theory (1736–1936).* Oxford, England: Clarendon Press, 1976.
10. Bondy, J. A., and Murty, U. S. R. *Graph Theory with Applications.* New York: Elsevier North-Holland, 1976.
11. Buckley, Fred, and Harary, Frank. *Distance in Graphs.* Reading, Mass.: Addison-Wesley, 1990.
12. Chartrand, Gary, and Lesniak. Linda. *Graphs and Digraphs,* 3rd ed. Boca Raton, Fla.: CRC Press, 1996.
13. Chartrand, Gary, and Wilson, Robin J."The Petersen Graph." In Frank Harary and John S. Maybee, eds., *Graphs and Applications.* New York: Wiley, 1985.

14. Deo, Narsingh. *Graph Theory with Applications to Engineering and Computer Science.* Englewood Cliffs, N. J.: Prentice-Hall, 1974.
15. Elsayed, E. A."Algorithms for Optimal Material Handling in Automatic Warehousing Systems." *Int. J. Prod. Res.* 19 (1981): pp. 525–535.
16. Elsayed, E. A., and Stern, R. G."Computerized Algorithms for Order Processing in Automated Warehousing Systems." *Int. J. Prod. Res.* 21 (1983): pp. 579–586.
17. Foulds, L. R. *Combinatorial Optimization for Undergraduates.* New York: Springer-Verlag, 1984.
18. Harary, Frank. *Graph Theory.* Reading, Mass.: Addison-Wesley, 1969.
19. Harary, Frank, and Palmer, Edgar M. *Graphical Enumeration.* New York: Academic Press, 1973.
20. Hartsfield, Nora, and Ringel, Gerhard. *Pearls in Graph Theory: A Comprehensive Introduction.* Boston, Mass.: Harcourt/Academic Press, 1994.
21. Jünger, M., Reinelt, G., and Rinaldi, G."The Traveling Salesman Problem." In M. O. Ball, Thomas L. Magnanti, C. L. Monma, and G. L. Nemhauser, eds., *Handbooks in Operations Research and Management Science*, Vol. 7, *Network Models.* Amsterdam, Holland: Elsevier, 1995, pp. 225–330.
22. Lawler, E. L., Lenstra, J. K., Rinnooy Kan, A. H. G., and Shmoys, D. B., eds. *The Traveling Salesman Problem.* New York: Wiley, 1986.
23. Liu, C. L. *Introduction to Combinatorial Mathematics.* New York: McGraw-Hill, 1968.
24. Ralston, Anthony. "De Bruijn Sequences — A Model Example of the Interaction of Discrete Mathematics and Computer Science." *Mathematics Magazine* 55, no. 3 (May 1982): pp. 131–143.
25. Read, R. C."An Introduction to Chromatic Polynomials." *Journal of Combinatorial Theory* 4 (1968): pp. 52–71.
26. Roberts, Fred S. *Discrete Mathematical Models.* Englewood Cliffs, N. J.: Prentice-Hall, 1976.
27. Robertson, N., Sanders, D. P., Seymour, P. D., and Thomas, R. "Efficiently Four-coloring Planar Graphs." *Proceedings of the 28th ACM Symposium on the Theory of Computation.* ACM Press (1996): pp. 571–575.
28. Robertson, N., Sanders, D. P., Seymour, P. D., and Thomas, R. "The Four-color Theorem." *Journal of Combinatorial Theory Series B* 70 (1997): pp. 166–183.
29. Rouvray, Dennis H. "Predicting Chemistry from Topology." *Scientific American* 255, no. 3 (September 1986): pp. 40–47.
30. Seshu, S., and Reed, M. B. *Linear Graphs and Electrical Networks.* Reading, Mass.: Addison-Wesley, 1961.
31. Tymoczko, Thomas. "Computers, Proofs and Mathematicians: A Philosophical Investigation of the Four-Color Proof." *Mathematics Magazine* 53, no. 3 (May 1980): pp. 131–138.
32. West, Douglas B. *Introduction to Graph Theory*, 2nd ed. Upper Saddle River, N.J.: Prentice-Hall, 2001.

SUPPLEMENTARY EXERCISES

1. Let G be a loop-free undirected graph on n vertices. If G has 56 edges and $\overline{G}$ has 80 edges, what is n?

2. Determine the number of cycles of length 4 in the hypercube Q_n.

3. **a)** If the edges of K_6 are painted either red or blue, prove that there is a red triangle or a blue triangle that is a subgraph.

b) Prove that in any group of six people there must be three who are total strangers to one another or three who are mutual friends.

4. **a)** Let $G = (V, E)$ be a loop-free undirected graph. Recall that G is called self-complementary if G and $\overline{G}$ are isomorphic. If G is self-complementary (i) determine $|E|$ if $|V| = n$; (ii) prove that G is connected.

b) Let $n \in \mathbf{Z}^+$, where $n = 4k$ $(k \in \mathbf{Z}^+)$ or $n = 4k + 1$ $(k \in \mathbf{N})$. Prove that there exists a self-complementary graph $G = (V, E)$, where $|V| = n$.

5. **a)** Show that the graphs G_1 and G_2, in Fig. 11.95, are isomorphic.

b) How many different isomorphisms $f: G_1 \to G_2$ are possible here?

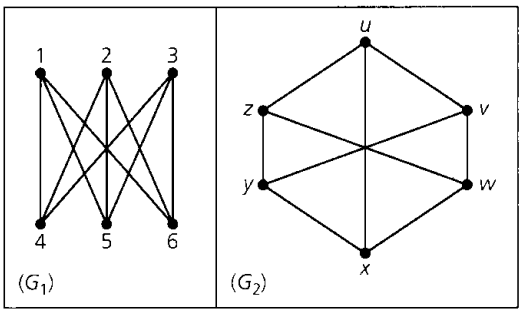


Figure 11.95

6. Are any of the planar graphs for the five Platonic solids bipartite?

7. **a)** How many paths of length 5 are there in the complete bipartite graph $K_{3,7}$? (Remember that a path such as $v_1 \to v_2 \to v_3 \to v_4 \to v_5 \to v_6$ is considered to be the same as the path $v_6 \to v_5 \to v_4 \to v_3 \to v_2 \to v_1$.)

b) How many paths of length 4 are there in $K_{3,7}$?

c) Let $m, n, p \in \mathbf{Z}^+$ with $2m < n$ and $1 \leq p \leq 2m$. How many paths of length p are there in the complete bipartite graph $K_{m,n}$?

8. Let $X = \{1, 2, 3, \ldots, n\}$, where $n \geq 2$. Construct the loop-free undirected graph $G = (V, E)$ as follows:

- (V): Each two-element subset of X determines a vertex of G.
- (E): If $v_1, v_2 \in V$ correspond to subsets $\{a, b\}$ and $\{c, d\}$, respectively, of X, draw the edge $\{v_1, v_2\}$ in G when $\{a, b\} \cap \{c, d\} = \emptyset$.

a) Show that G is an isolated vertex when $n = 2$ and that G is disconnected for $n = 3, 4$.

b) Show that for $n \geq 5$, G is connected. (In fact, for all $v_1, v_2 \in V$, either $\{v_1, v_2\} \in E$ or there is a path of length 2 connecting v_1 and v_2.)

c) Prove that G is nonplanar for $n \geq 5$.

d) Prove that for $n \geq 8$, G has a Hamilton cycle.

9. If $G = (V, E)$ is an undirected graph, a subset K of V is called a *covering* of G if for every edge $\{a, b\}$ of G either a or b is in K. The set K is a *minimal covering* if $K - \{x\}$ fails to cover G for each $x \in K$. The number of vertices in a smallest covering is called the *covering number* of G.

a) Prove that if $I \subseteq V$, then I is an independent set in G if and only if $V - I$ is a covering of G.

b) Verify that $|V|$ is the sum of the independence number of G (as defined in Exercise 25 for Section 11.5) and its covering number.

10. If $G = (V, E)$ is an undirected graph, a subset D of V is called a *dominating set* if for all $v \in V$, either $v \in D$ or v is adjacent to a vertex in D. If D is a dominating set and no proper subset of D has this property, then D is called *minimal*. The size of any smallest dominating set in G is denoted by $\gamma(G)$ and is called the *domination number* of G.

a) If G has no isolated vertices, prove that if D is a minimal dominating set, then $V - D$ is a dominating set.

b) If $I \subseteq V$ is independent, prove that I is a dominating set if and only if I is maximal independent.

c) Show that $\gamma(G) \leq \beta(G)$, and that $|V| \leq \beta(G)\chi(G)$. [Here $\beta(G)$ is the independence number of G — first given in Exercise 25 of Section 11.5.]

11. Let $G = (V, E)$ be the undirected connected "ladder graph" shown in Fig. 11.94. For $n \geq 0$, let a_n denote the number of ways one can select n of the edges in G so that no two edges share a common vertex. Find and solve a recurrence relation for a_n.

12. Consider the four *comb* graphs in parts (i), (ii), (iii), and (iv) of Fig. 11.96. These graphs have 1 tooth, 2 teeth, 3 teeth, and n teeth, respectively. For $n \geq 1$, let a_n count the number of independent subsets in $\{x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n\}$. Find and solve a recurrence relation for a_n.

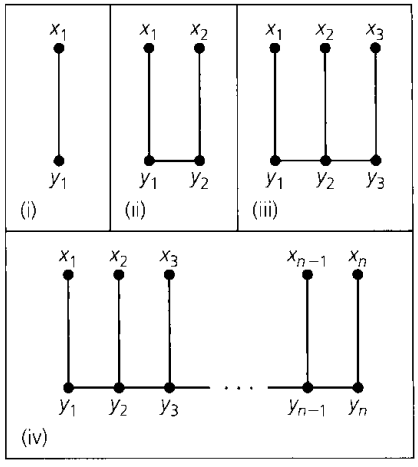


Figure 11.96

13. Consider the four graphs in parts (i), (ii), (iii), and (iv) of Fig. 11.97. If a_n counts the number of independent subsets of $\{x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n\}$, where $n \geq 1$, find and solve a recurrence relation for a_n.

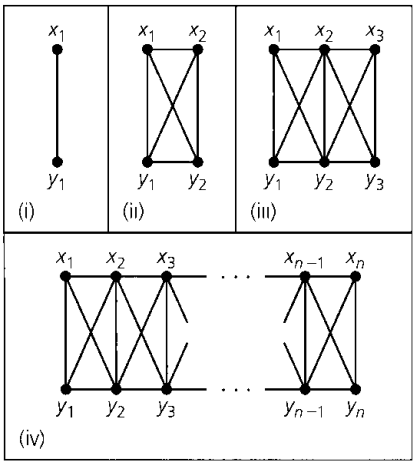


Figure 11.97

14. For $n \geq 1$, let $a_n = \binom{n}{2}$, the number of edges in K_n, and let $a_0 = 0$. Find the generating function $f(x) = \sum_{n=0}^{\infty} a_n x^n$.

15. For the graph G in Fig. 11.98, answer the following questions.

a) What are $\gamma(G)$, $\beta(G)$, and $\chi(G)$?

b) Does G have an Euler circuit or a Hamilton cycle?

c) Is G bipartite? Is it planar?

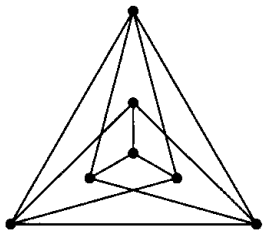


Figure 11.98

16. a) Suppose that the complete bipartite graph $K_{m,n}$ contains 16 edges and satisfies $m \leq n$. Determine m, n so that $K_{m,n}$ possesses (i) an Euler circuit but not a Hamilton cycle; (ii) both a Hamilton cycle and an Euler circuit.

b) Generalize the results of part (a).

17. If $G = (V, E)$ is an undirected graph, any subgraph of G that is a complete graph is called a *clique* in G. The number of vertices in a largest clique in G is called the *clique number* for G and is denoted by $\omega(G)$.

a) How are $\chi(G)$ and $\omega(G)$ related?

b) Is there any relationship between $\omega(G)$ and $\beta(\overline{G})$?

18. If $G = (V, E)$ is an undirected loop-free graph, the *line graph* of G, denoted $L(G)$, is a graph with the set E as vertices, where we join two vertices e_1, e_2 in $L(G)$ if and only if e_1, e_2 are adjacent edges in G.

a) Find $L(G)$ for each of the graphs in Fig. 11.99.

b) Assuming that $|V| = n$ and $|E| = e$, show that $L(G)$ has e vertices and $(1/2)\sum_{v \in V} \deg(v)[\deg(v) - 1] = [(1/2)\sum_{v \in V}[\deg(v)]^2] - e = \sum_{v \in V} \binom{\deg(v)}{2}$ edges.

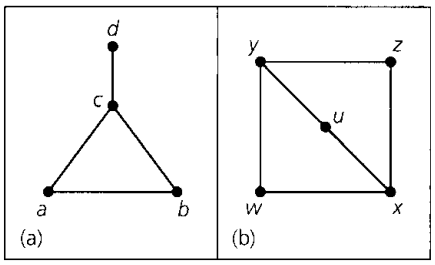


Figure 11.99

c) Prove that if G has an Euler circuit, then $L(G)$ has both an Euler circuit and a Hamilton cycle.

d) If $G = K_4$, examine $L(G)$ to show that the converse of part (c) is false.

e) Prove that if G has a Hamilton cycle, then so does $L(G)$.

f) Examine $L(G)$ for the graph in Fig. 11.99(b) to show that the converse of part (e) is false.

g) Verify that $L(G)$ is nonplanar for $G = K_5$ and $G = K_{3,3}$.

h) Give an example of a graph G, where G is planar but $L(G)$ is not.

19. Explain why each of the following polynomials in λ cannot be a chromatic polynomial.

a) $\lambda^4 - 5\lambda^3 + 7\lambda^2 - 6\lambda + 3$

b) $3\lambda^3 - 4\lambda^2 + \lambda$

c) $\lambda^4 - 3\lambda^3 + 5\lambda^2 - 4\lambda$

20. a) For all $x, y \in \mathbf{Z}^+$, prove that $x^3y - xy^3$ is even.

b) Let $V = \{1, 2, 3, \ldots, 8, 9\}$. Construct the loop-free undirected graph $G = (V, E)$ as follows: For $m, n \in V$, $m \neq n$, draw the edge $\{m, n\}$ in G if 5 divides $m + n$ or $m - n$.

c) Given any three distinct positive integers, prove that there are two of these, say x and y, where 10 divides $x^3y - xy^3$.

21. a) For $n \geq 1$, let P_{n-1} denote the path made up of n vertices and $n - 1$ edges. Let a_n be the number of independent subsets of vertices in P_{n-1}. (The empty subset is considered one of these independent subsets.) Find and solve a recurrence relation for a_n.

b) Determine the number of independent subsets (of vertices) in each of the graphs G_1, G_2, and G_3, of Fig. 11.100.

c) For each of the graphs H_1, H_2, and H_3, of Fig. 11.101, find the number of independent subsets of vertices.

d) Let $G = (V, E)$ be a loop-free undirected graph with $V = \{v_1, v_2, \ldots, v_r\}$ and where there are m independent subsets of vertices. The graph $G' = (V', E')$ is constructed from G as follows: $V' = V \cup \{x_1, x_2, \ldots, x_s\}$, with no x_i in V, for all $1 \leq i \leq s$; and $E' = E \cup \{\{x_i, v_j\} | 1 \leq i \leq s, 1 \leq j \leq r\}$. How many subsets of V' are independent?

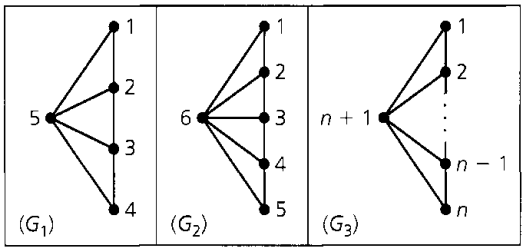


Figure 11.100

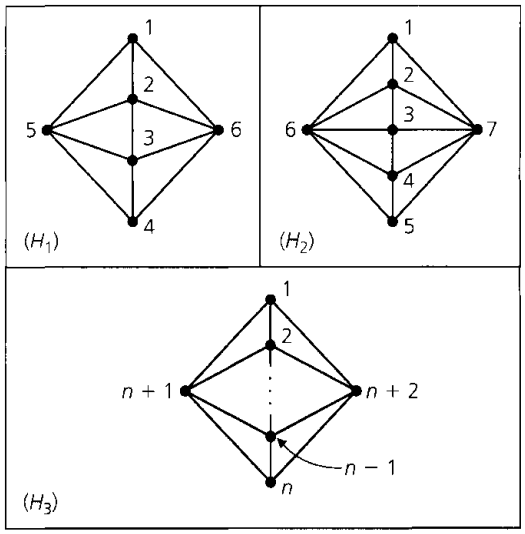


Figure 11.101

22. Suppose that $G = (V, E)$ is a loop-free undirected graph. If G is 5-regular and $|V| = 10$, prove that G is nonplanar.

12

Trees

Continuing our study of graph theory, we shall now focus on a special type of graph called a tree. First used in 1847 by Gustav Kirchhoff (1824–1887) in his work on electrical networks, trees were later redeveloped and named by Arthur Cayley (1821–1895). In 1857 Cayley used these special graphs in order to enumerate the different isomers of the saturated hydrocarbons C_nH_{2n+2}, $n \in \mathbf{Z}^+$.

With the advent of digital computers, many new applications were found for trees. Special types of trees are prominent in the study of data structures, sorting, and coding theory, and in the solution of certain optimization problems.

12.1 Definitions, Properties, and Examples

Definition 12.1 Let $G = (V, E)$ be a loop-free undirected graph. The graph G is called a *tree*[†] if G is connected and contains no cycles.

In Fig. 12.1 the graph G_1 is a tree, but the graph G_2 is not a tree because it contains the cycle $\{a, b\}, \{b, c\}, \{c, a\}$. The graph G_3 is not connected, so it cannot be a tree. However, each component of G_3 is a tree, and in this case we call G_3 a *forest.*

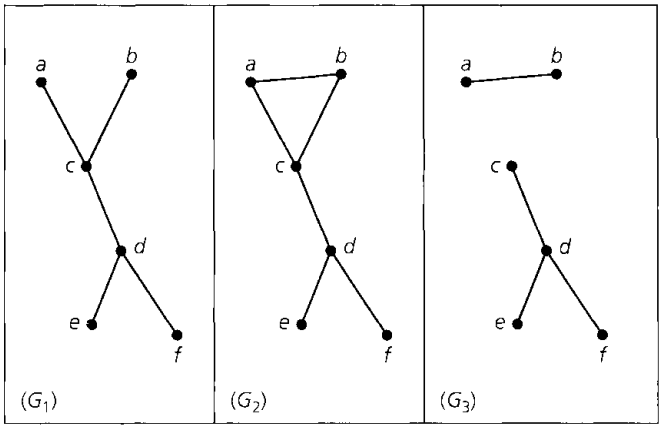


Figure 12.1

[†]As in the case of graphs, the terminology in the study of trees is not standard and the reader may find some differences from one textbook to another.

When a graph is a tree we write T instead of G to emphasize this structure.

In Fig. 12.1 we see that G_1 is a subgraph of G_2 where G_1 contains all the vertices of G_2 and G_1 is a tree. In this situation G_1 is a spanning tree for G_2. Hence a *spanning tree* for a connected graph is a spanning subgraph that is also a tree. We may think of a spanning tree as providing minimal connectivity for the graph and as a minimal skeletal framework holding the vertices together. The graph G_3 provides a *spanning forest* for the graph G_2.

We now examine some properties of trees.

THEOREM 12.1

If a, b are distinct vertices in a tree $T = (V, E)$, then there is a unique path that connects these vertices.

Proof: Since T is connected, there is at least one path in T that connects a and b. If there were more, then from two such paths some of the edges would form a cycle. But T has no cycles.

THEOREM 12.2

If $G = (V, E)$ is an undirected graph, then G is connected if and only if G has a spanning tree.

Proof: If G has a spanning tree T, then for every pair a, b of distinct vertices in V a subset of the edges in T provides a (unique) path between a and b, and so G is connected. Conversely, if G is connected and G is not a tree, remove all loops from G. If the resulting subgraph G_1 is not a tree, then G_1 must contain a cycle C_1. Remove an edge e_1 from C_1 and let $G_2 = G_1 - e_1$. If G_2 contains no cycles, then G_2 is a spanning tree for G because G_2 contains all the vertices in G, is loop-free, and is connected. If G_2 does contain a cycle — say, C_2 — then remove an edge e_2 from C_2 and consider the subgraph $G_3 = G_2 - e_2 = G_1 - \{e_1, e_2\}$. Once again, if G_3 contains no cycles, then we have a spanning tree for G. Otherwise we continue this procedure a finite number of additional times until we arrive at a spanning subgraph of G that is loop-free and connected and contains no cycles (and, consequently, is a spanning tree for G).

Figure 12.2 shows three nonisomorphic trees that exist for five vertices. Although they are not isomorphic, they all have the same number of edges, namely, four. This leads us to the following general result.

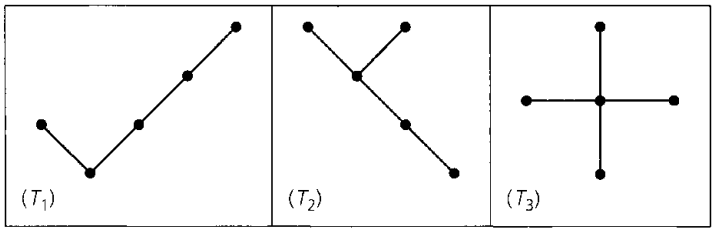


Figure 12.2

THEOREM 12.3

In every tree $T = (V, E)$, $|V| = |E| + 1$.

Proof: The proof is obtained by applying the alternative form of the Principle of Mathematical Induction to $|E|$. If $|E| = 0$, then the tree consists of a single isolated vertex, as in

Fig. 12.3(a). Here $|V| = 1 = |E| + 1$. Parts (b) and (c) of the figure verify the result for the cases where $|E| = 1$ or 2.

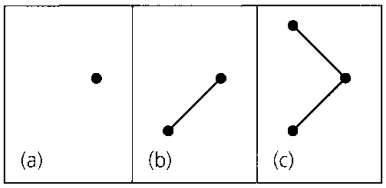


Figure 12.3

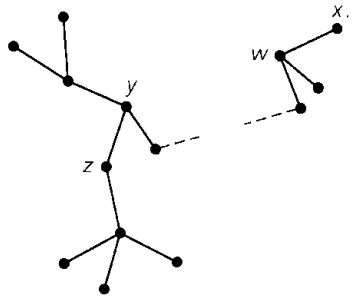


Figure 12.4

Assume the theorem is true for every tree that contains at most k edges, where $k \geq 0$. Now consider a tree $T = (V, E)$, as in Fig. 12.4, where $|E| = k + 1$. [The dotted edge(s) indicates that some of the tree doesn't appear in the figure.] If, for instance, the edge with endpoints y, z is removed from T, we obtain two *subtrees*, $T_1 = (V_1, E_1)$ and $T_2 = (V_2, E_2)$, where $|V| = |V_1| + |V_2|$ and $|E_1| + |E_2| + 1 = |E|$. (One of these subtrees could consist of just a single vertex if, for example, the edge with endpoints w, x were removed.) Since $0 \leq |E_1| \leq k$ and $0 \leq |E_2| \leq k$, it follows, by the induction hypothesis, that $|E_i| + 1 = |V_i|$, for $i = 1, 2$. Consequently, $|V| = |V_1| + |V_2| = (|E_1| + 1) + (|E_2| + 1) = (|E_1| + |E_2| + 1) + 1 = |E| + 1$, and the theorem follows by the alternative form of the Principle of Mathematical Induction.

As we examine the trees in Fig. 12.2 we also see that each tree has at least two pendant vertices — that is, vertices of degree 1. This is also true in general.

THEOREM 12.4

For every tree $T = (V, E)$, if $|V| \geq 2$, then T has at least two pendant vertices.

Proof: Let $|V| = n \geq 2$. From Theorem 12.3 we know that $|E| = n - 1$, so by Theorem 11.2 it follows that $2(n - 1) = 2|E| = \sum_{v \in V} \deg(v)$. Since T is connected, we have $\deg(v) \geq 1$ for all $v \in V$. If there are k pendant vertices in T, then each of the other $n - k$ vertices has degree at least 2 and

$$2(n - 1) = 2|E| = \sum_{v \in V} \deg(v) \geq k + 2(n - k).$$

From this we see that $[2(n - 1) \geq k + 2(n - k)] \Rightarrow [(2n - 2) \geq (k + 2n - 2k)] \Rightarrow [-2 \geq -k] \Rightarrow [k \geq 2]$, and the result is consequently established.

EXAMPLE 12.1

In Fig. 12.5 we have two trees, each with 14 vertices (labeled with C's and H's) and 13 edges. Each vertex has degree 4 (C, carbon atom) or degree 1 (H, hydrogen atom). Part (b) of the figure has a carbon atom (C) at the center of the tree. This carbon atom is adjacent to four vertices, three of which have degree 4. There is no vertex (C atom) in part (a) that possesses this property, so the two trees are not isomorphic. They serve as models for the two chemical

isomers that correspond with the saturated† hydrocarbon C_4H_{10}. Part (a) represents n-butane (formerly called butane); part (b) represents 2-methyl propane (formerly called isobutane).

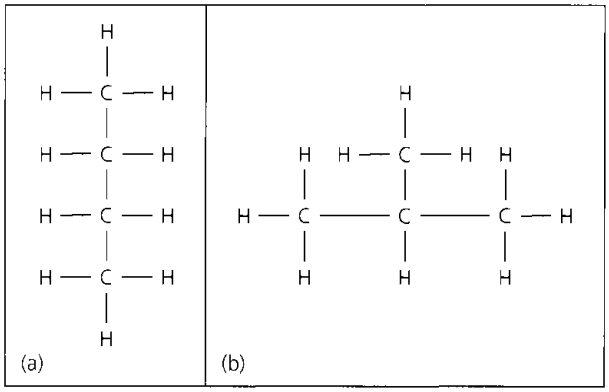


Figure 12.5

A second result from chemistry is given in the following example.

EXAMPLE 12.2

If a saturated hydrocarbon [in particular, an acyclic (no cycles), single-bond hydrocarbon—called an *alkane*] has n carbon atoms, show that it has $2n + 2$ hydrogen atoms.

Considering the saturated hydrocarbon as a tree $T = (V, E)$, let k equal the number of pendant vertices, or hydrogen atoms, in the tree. Then with a total of $n + k$ vertices, where each of the n carbon atoms has degree 4, we find that

$$4n + k = \sum_{v \in V} \deg(v) = 2|E| = 2(|V| - 1) = 2(n + k - 1),$$

and

$$4n + k = 2(n + k - 1) \Rightarrow k = 2n + 2.$$

We close this section with a theorem that provides several different ways to characterize trees.

THEOREM 12.5

The following statements are equivalent for a loop-free undirected graph $G = (V, E)$.

a) G is a tree.

b) G is connected, but the removal of any edge from G disconnects G into two subgraphs that are trees.

c) G contains no cycles, and $|V| = |E| + 1$.

d) G is connected, and $|V| = |E| + 1$.

†The adjective *saturated* is used here to indicate that for the number of carbon atoms present in the molecule, we have the maximum number of hydrogen atoms.

e) G contains no cycles, and if $a, b \in V$ with $\{a, b\} \notin E$, then the graph obtained by adding edge $\{a, b\}$ to G has precisely one cycle.

Proof: We shall prove that (a) $\Rightarrow$ (b), (b) $\Rightarrow$ (c), and (c) $\Rightarrow$ (d), leaving to the reader the proofs for (d) $\Rightarrow$ (e) and (e) $\Rightarrow$ (a).

[(a) $\Rightarrow$ (b)]: If G is a tree, then G is connected. So let $e = \{a, b\}$ be any edge of G. Then if $G - e$ is connected, there are at least two paths in G from a to b. But this contradicts Theorem 12.1. Hence $G - e$ is disconnected and so the vertices in $G - e$ may be partitioned into two subsets: (1) vertex a and those vertices that can be reached from a by a path in $G - e$; and (2) vertex b and those vertices that can be reached from b by a path in $G - e$. These two connected components are trees because a loop or cycle in either component would also be in G.

[(b) $\Rightarrow$ (c)]: If G contains a cycle, then let $e = \{a, b\}$ be an edge of the cycle. But then $G - e$ is connected, contradicting the hypothesis in part (b). So G contains no cycles, and since G is a loop-free connected undirected graph, we know that G is a tree. Consequently, it follows from Theorem 12.3 that $|V| = |E| + 1$.

[(c) $\Rightarrow$ (d)]: Let $\kappa(G) = r$ and let $G_1, G_2, \ldots, G_r$ be the components of G. For $1 \leq i \leq r$, select a vertex $v_i \in G_i$ and add the $r - 1$ edges $\{v_1, v_2\}, \{v_2, v_3\}, \ldots, \{v_{r-1}, v_r\}$ to G to form the graph $G' = (V, E')$, which is a tree. Since G' is a tree, we know that $|V| = |E'| + 1$ because of Theorem 12.3. But from part (c), $|V| = |E| + 1$, so $|E| = |E'|$ and $r - 1 = 0$. With $r = 1$, it follows that G is connected.

EXERCISES 12.1

1. a) Draw the graphs of all nonisomorphic trees on six vertices.

b) How many isomers does hexane (C_6H_{14}) have?

2. Let $T_1 = (V_1, E_1)$, $T_2 = (V_2, E_2)$ be two trees where $|E_1| = 17$ and $|V_2| = 2|V_1|$. Determine $|V_1|$, $|V_2|$, and $|E_2|$.

3. a) Let $F_1 = (V_1, E_1)$ be a forest of seven trees where $|E_1| = 40$. What is $|V_1|$?

b) If $F_2 = (V_2, E_2)$ is a forest with $|V_2| = 62$ and $|E_2| = 51$, how many trees determine F_2?

4. If $G = (V, E)$ is a forest with $|V| = v$, $|E| = e$, and κ components (trees), what relationship exists among v, e, and κ?

5. What kind of trees have exactly two pendant vertices?

6. a) Verify that all trees are planar.

b) Derive Theorem 12.3 from part (a) and Euler's Theorem for planar graphs.

7. Give an example of an undirected graph $G = (V, E)$ where $|V| = |E| + 1$ but G is not a tree.

8. a) If a tree has four vertices of degree 2, one vertex of degree 3, two of degree 4, and one of degree 5, how many pendant vertices does it have?

b) If a tree $T = (V, E)$ has v_2 vertices of degree 2, v_3 vertices of degree 3, ..., and v_m vertices of degree m, what are $|V|$ and $|E|$?

9. If $G = (V, E)$ is a loop-free undirected graph, prove that G is a tree if there is a unique path between any two vertices of G.

10. The connected undirected graph $G = (V, E)$ has 30 edges. What is the maximum value that $|V|$ can have?

11. Let $T = (V, E)$ be a tree with $|V| = n \geq 2$. How many distinct paths are there (as subgraphs) in T?

12. Let $G = (V, E)$ be a loop-free connected undirected graph where $V = \{v_1, v_2, v_3, \ldots, v_n\}$, $n \geq 2$, $\deg(v_1) = 1$, and $\deg(v_i) \geq 2$ for $2 \leq i \leq n$. Prove that G must have a cycle.

13. Find two nonisomorphic spanning trees for the complete bipartite graph $K_{2,3}$. How many nonisomorphic spanning trees are there for $K_{2,3}$?

14. For $n \in \mathbf{Z}^+$, how many nonisomorphic spanning trees are there for $K_{2,n}$?

15. Determine the number of nonidentical (though some may be isomorphic) spanning trees that exist for each of the graphs shown in Fig. 12.6.

16. For each graph in Fig. 12.7, determine how many nonidentical (though some may be isomorphic) spanning trees exist.

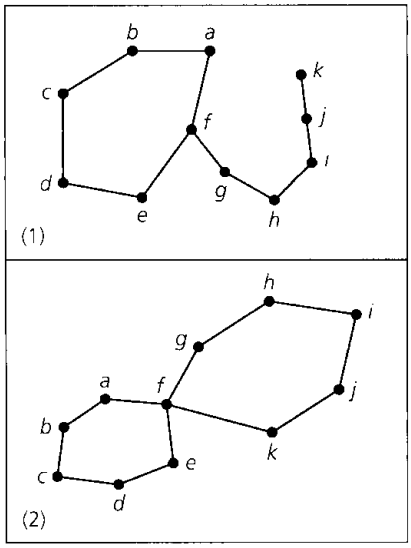


Figure 12.6

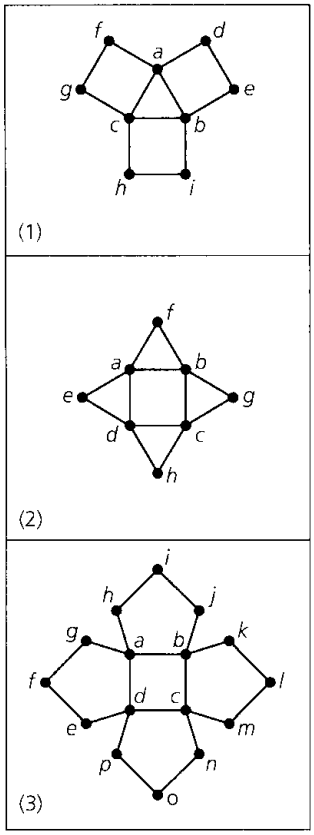


Figure 12.7

17. Let $T = (V, E)$ be a tree where $|V| = n$. Suppose that for each $v \in V$, $\deg(v) = 1$ or $\deg(v) \geq m$, where m is a fixed positive integer and $m \geq 2$.

a) What is the smallest value possible for n?

b) Prove that T has at least m pendant vertices.

18. Suppose that $T = (V, E)$ is a tree with $|V| = 1000$. What is the sum of the degrees of all the vertices in T?

19. Let $G = (V, E)$ be a loop-free connected undirected graph. Let H be a subgraph of G. The *complement of H in G* is the subgraph of G made up of those edges in G that are not in H (along with the vertices incident to these edges).

a) If T is a spanning tree of G, prove that the complement of T in G does not contain a cut-set of G.

b) If C is a cut-set of G, prove that the complement of C in G does not contain a spanning tree of G.

20. Complete the proof of Theorem 12.5.

21. A labeled tree is one wherein the vertices are labeled. If the tree has n vertices, then $\{1, 2, 3, \ldots, n\}$ is used as the set of labels. We find that two trees that are isomorphic without labels may become nonisomorphic when labeled. In Fig. 12.8, the first two trees are isomorphic as labeled trees. The third tree is isomorphic to the other two if we ignore the labels; as a labeled tree, however, it is not isomorphic to either of the other two.

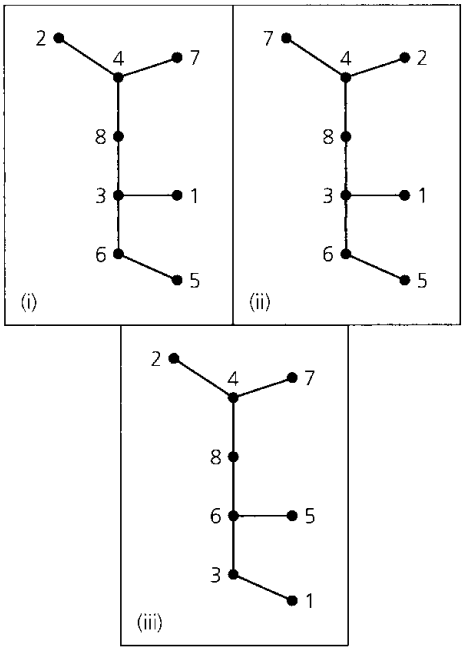


Figure 12.8

The number of nonisomorphic trees with n labeled vertices can be counted by setting up a one-to-one correspondence between these trees and the n^{n-2} sequences (with repetitions allowed) $x_1, x_2, \ldots, x_{n-2}$ whose entries are taken from $\{1, 2, 3, \ldots, n\}$. If T is one such labeled tree, we use the following algorithm to find its corresponding sequence — called the *Prüfer code* for the tree. (Here T has at least one edge.)

Step 1: Set the counter i to 1.

Step 2: Set $T(i) = T$.

Step 3: Since a tree has at least two pendant vertices, select the pendant vertex in $T(i)$ with the smallest label y_i. Now remove the edge $\{x_i, y_i\}$ from $T(i)$ and use x_i for the ith component of the sequence.

Step 4: If $i = n - 2$, we have the sequence corresponding to the given labeled tree $T(1)$. If $i \neq n - 2$, increase i by 1, set $T(i)$ equal to the resulting subtree obtained in step (3), and return to step (3).

a) Find the six-digit sequence (Prüfer code) for trees (i) and (iii) in Fig. 12.8.

b) If v is a vertex in T, show that the number of times the label on v appears in the Prüfer code $x_1, x_2, \ldots, x_{n-2}$ is $\deg(v) - 1$.

c) Reconstruct the labeled tree on eight vertices that is associated with the Prüfer code 2, 6, 5, 5, 5, 5.

d) Develop an algorithm for reconstructing a tree from a given Prüfer code $x_1, x_2, \ldots, x_{n-2}$.

22. Let $n \in \mathbf{Z}^+$, $n \geq 3$. If v is a vertex in K_n, how many of the n^{n-2} spanning trees of K_n have v as a pendant vertex?

23. Characterize the trees whose Prüfer codes

a) contain only one integer, or

b) have distinct integers in all positions.

24. Show that the number of labeled trees with n vertices, k of which are pendant vertices, is $\binom{n}{k}(n-k)!S(n-2, n-k) = (n!/k!)S(n-2, n-k)$, where $S(n-2, n-k)$ is a Stirling number of the second kind. (This result was first established in 1959 by A. Rényi.)

25. Let $G = (V, E)$ be the undirected graph in Fig. 12.9. Show that the edge set E can be partitioned as $E_1 \cup E_2$ so that the subgraphs $G_1 = (V, E_1)$, $G_2 = (V, E_2)$ are isomorphic spanning trees of G.

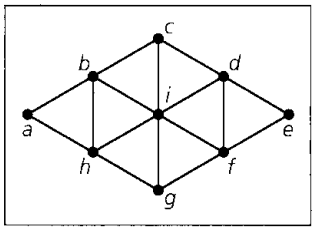


Figure 12.9

12.2 Rooted Trees

We turn now to directed trees. We find a variety of applications for a special type of directed tree called a rooted tree.

Definition 12.2 If G is a directed graph, then G is called a *directed tree* if the undirected graph associated with G is a tree. When G is a directed tree, G is called a *rooted tree* if there is a unique vertex r, called the *root*, in G with the in degree of $r = id(r) = 0$, and for all other vertices v, the in degree of $v = id(v) = 1$.

The tree in part (a) of Fig. 12.10 is directed but not rooted; the tree in part (b) is rooted with root r.

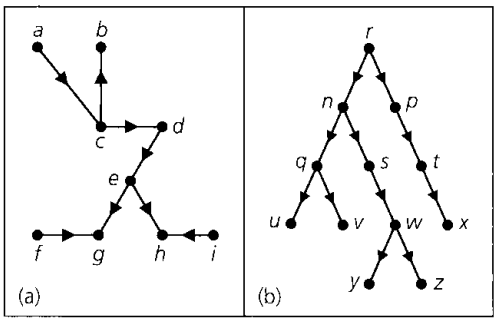


Figure 12.10

We draw rooted trees as in Fig. 12.10(b) but with the directions understood as going from the upper level to the lower level, so that the arrows aren't needed. In a rooted tree, a vertex with out degree 0 is called a *leaf* (or *terminal vertex.*) Vertices u, v, x, y, z are leaves in Fig. 12.10(b). All other vertices are called *branch nodes* (or *internal vertices*).

Consider the vertex s in this rooted tree [Fig. 12.10(b)]. The path from the root, r, to s is of length 2, so we say that s is at *level* 2 in the tree, or that s has *level number* 2. Similarly, x is at level 3, whereas y has level number 4. We call s a *child* of n, and we call n the *parent* of s. Vertices w, y, and z are considered *descendants* of s, n, and r, while s, n, and r are called *ancestors* of w, y, and z. In general, if v_1 and v_2 are vertices in a rooted tree and v_1 has the smaller level number, then v_1 is an ancestor of v_2 (or v_2 is a descendant of v_1) if there is a (directed) path from v_1 to v_2. Two vertices with a common parent are referred to as *siblings*. Such is the case for vertices q and s, whose common parent is vertex n. Finally, if v_1 is any vertex of the tree, the *subtree at* v_1 is the subgraph induced by the root v_1 and all of its descendants (there may be none).

EXAMPLE 12.3

In Fig. 12.11(a) a rooted tree is used to represent the table of contents of a three-chapter ($C1$, $C2$, $C3$) book. Vertices with level number 2 are for sections within a chapter; those at level 3 represent subsections within a section. Part (b) of the figure displays the natural order for the table of contents of this book.

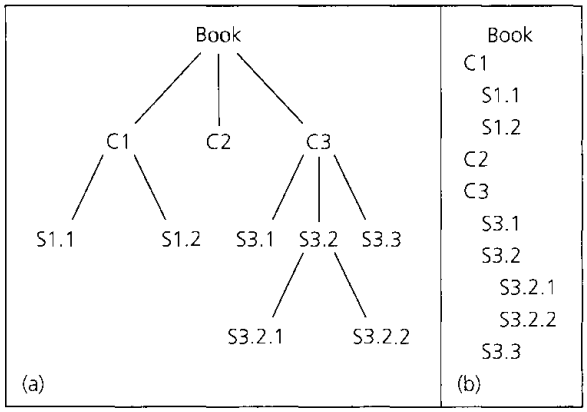


Figure 12.11

The tree in Fig. 12.11(a) suggests an order for the vertices if we examine the subtrees at $C1$, $C2$, and $C3$ from left to right. (This order will recur again in this section, in a more general context.) We now consider a second example that provides such an order.

EXAMPLE 12.4

In the tree T shown in Fig. 12.12, the edges (or branches, as they are often called) leaving each internal vertex are *ordered* from left to right. Hence T is called an *ordered rooted tree.*

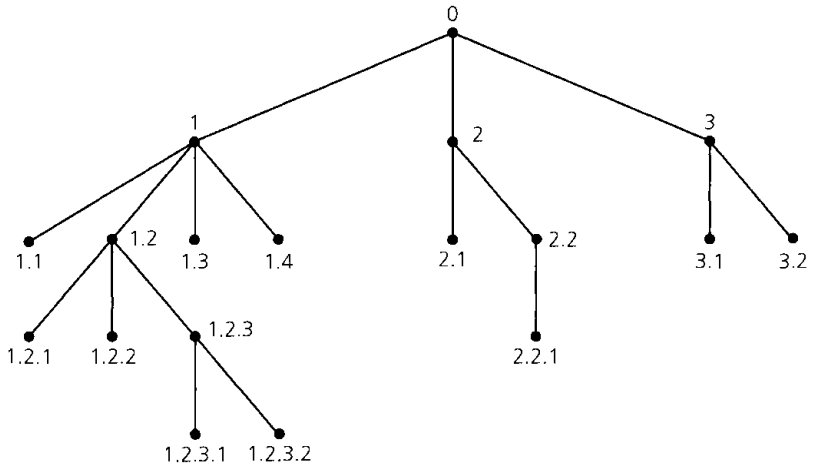


Figure 12.12

We label the vertices for this tree by the following algorithm.

Step 1: First assign the root the label (or *address*) 0.

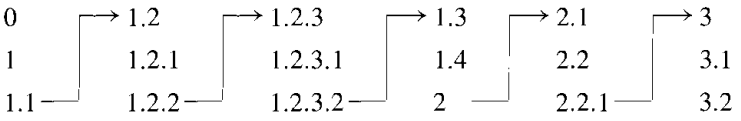
Step 2: Next assign the positive integers 1, 2, 3, . . . to the vertices at level 1, going from left to right.

Step 3: Now let v be an internal vertex at level $n \geq 1$, and let $v_1, v_2, \ldots, v_k$ denote the children of v (going from left to right). If a is the label assigned to vertex v, assign the labels $a.1, a.2, \ldots, a.k$ to the children $v_1, v_2, \ldots, v_k$, respectively.

Consequently, each vertex in T, other than the root, has a label of the form $a_1.a_2.a_3 \ldots . a_n$ if and only if that vertex has level number n. This is known as the *universal address system.*

This system provides a way to *order* all vertices in T. If u and v are two vertices in T with addresses b and c, respectively, we define $b < c$ if (a) $b = a_1.a_2 \ldots . a_m$ and $c = a_1.a_2 \ldots . a_m.a_{m+1} \ldots . a_n$, with $m < n$; or (b) $b = a_1.a_2 \ldots . a_m.x_1 \ldots . y$ and $c = a_1.a_2 \ldots . a_m.x_2 \ldots . z$, where $x_1, x_2 \in \mathbf{Z}^+$ and $x_1 < x_2$.

For the tree under consideration, this ordering yields



Since this resembles the alphabetical ordering in a dictionary, the order is called the *lexicographic*, or *dictionary, order.*

We now consider an application of a rooted tree in the study of computer science.

EXAMPLE 12.5

a) A rooted tree is a *binary* rooted tree if for each vertex v, $od(v) = 0$, 1, or 2—that is, if v has at most two children. If $od(v) = 0$ or 2 for all $v \in V$, then the rooted tree is called a *complete* binary tree. Such a tree can represent a binary operation, as in parts

(a) and (b) of Fig. 12.13. To avoid confusion when dealing with a noncommutative operation $\circ$, we label the root as $\circ$ and require the result to be $a \circ b$, where a is the left child, and b the right child, of the root.

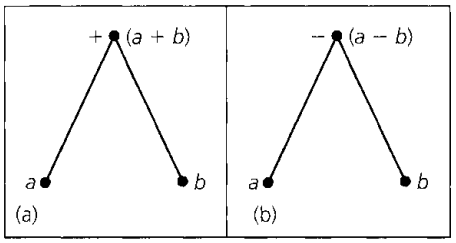


Figure 12.13

b) In Fig. 12.14 we extend the ideas presented in Fig. 12.13 in order to construct the binary rooted tree for the algebraic expression

$$((7 - a)/5) * ((a + b) \uparrow 3),$$

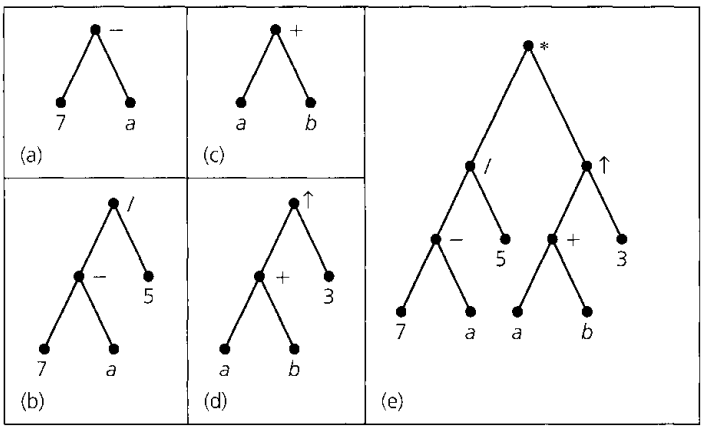


Figure 12.14

where $*$ denotes multiplication and $\uparrow$ denotes exponentiation. Here we construct this tree, as shown in part (e) of the figure, from the bottom up. First, a subtree for the expression $7 - a$ is constructed in part (a) of Fig. 12.14. This is then incorporated (as the left subtree for /) in the binary rooted tree for the expression $(7 - a)/5$ in Fig. 12.14 (b). Then, in a similar way, the binary rooted trees in parts (c) and (d) of the figure are constructed for the expressions $a + b$ and $(a + b) \uparrow 3$, respectively. Finally, the two subtrees in parts (b) and (d) are used as the left and right subtrees, respectively, for $*$ and give us the binary rooted tree [in Fig. 12.14(e)] for $((7 - a)/5) * ((a + b) \uparrow 3)$.

The same ideas are used in Fig. 12.15, where we find the binary rooted trees for the algebraic expressions

$$(a - (3/b)) + 5 \text{ [in part (a)]} \qquad \text{and} \qquad a - (3/(b + 5)) \text{ [in part (b)]}.$$

c) In evaluating $t + (uv)/(w + x - y^z)$ in certain procedural languages, we write the expression in the form $t + (u * v)/(w + x - y \uparrow z)$. When the computer evaluates this expression, it performs the binary operations (within each parenthesized part) according to a hierarchy of operations whereby exponentiation precedes multiplication

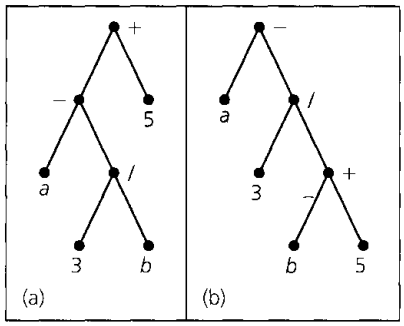


Figure 12.15

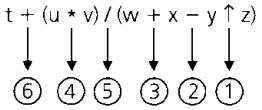


Figure 12.16

and division, which in turn precede addition and subtraction. In Fig. 12.16 we number the operations in the order in which they are performed by the computer. For the computer to evaluate this expression, it must somehow scan the expression in order to perform the operations in the order specified.

Instead of scanning back and forth continuously, however, the machine converts the expression into a notation that is independent of parentheses. This is known as Polish notation, in honor of the Polish (actually Ukrainian) logician Jan Lukasiewicz (1878–1956). Here the *infix* notation $a \circ b$ for a binary operation $\circ$ becomes $\circ\, ab$, the *prefix* (or Polish) notation. The advantage is that the expression in Fig. 12.16 can be rewritten without parentheses as

$$+\, t / * uv + w - x \uparrow yz,$$

where the evaluation proceeds from right to left. When a binary operation is encountered, it is performed on the two operands to its right. The result is then treated as one of the operands for the next binary operation encountered as we continue to the left. For instance, given the assignments $t = 4$, $u = 2$, $v = 3$, $w = 1$, $x = 9$, $y = 2$, $z = 3$, the following steps take place in the evaluation of the expression

$$+\, t / * uv + w - x \uparrow yz.$$

1) $+\ 4\ /\ *\ 2\ 3\ +\ 1\ -\ 9\ \underbrace{\uparrow\ 2\ 3}_{2 \uparrow 3 = 8}$

2) $+\ 4\ /\ *\ 2\ 3\ +\ 1\ \underbrace{-\ 9\ 8}_{9 - 8 = 1}$

3) $+\ 4\ /\ *\ 2\ 3\ \underbrace{+\ 1\ 1}_{1 + 1 = 2}$

4) $+\ 4\ /\ \underbrace{*\ 2\ 3}_{2 * 3 = 6}\ 2$

5) $+\ 4\ \underbrace{/\ 6\ 2}_{6\ /\ 2 = 3}$

6) $\underbrace{+\ 4\ 3}_{4 + 3 = 7}$

So the value of the given expression for the preceding assignments is 7.

The use of Polish notation is important for the compilation of computer programs and can be obtained by representing a given expression by a rooted tree, as shown in Fig. 12.17. Here each variable (or constant) is used to label a leaf of the tree. Each internal vertex is

labeled by a binary operation whose left and right operands are the left and right subtrees it determines. Starting at the root, as we transverse the tree from top to bottom and left to right, as shown in Fig. 12.17, we find the Polish notation by writing down the labels of the vertices in the order in which they are visited.

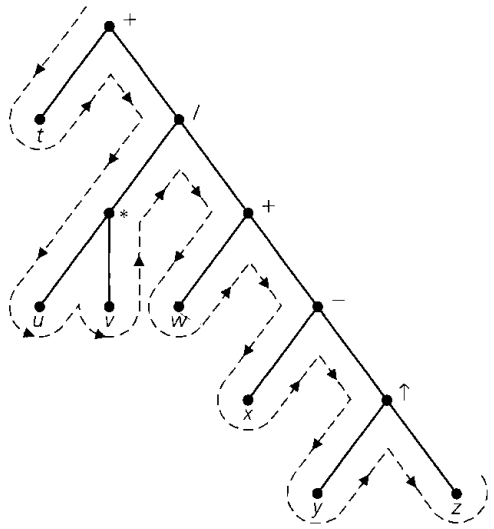


Figure 12.17

The last two examples illustrate the importance of order. Several methods exist for systematically ordering the vertices in a tree. Two of the most prevalent in the study of data structures are the preorder and postorder. These are defined recursively in the following definition.

Definition 12.3

Let $T = (V, E)$ be a rooted tree with root r. If T has no other vertices, then the root by itself constitutes the *preorder* and *postorder traversals* of T. If $|V| > 1$, let $T_1, T_2, T_3, \ldots, T_k$ denote the subtrees of T as we go from left to right (as in Fig. 12.18).

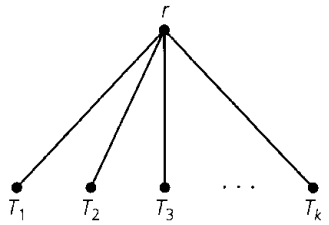


Figure 12.18

a) The *preorder traversal* of T first visits r and then traverses the vertices of T_1 in preorder, then the vertices of T_2 in preorder, and so on until the vertices of T_k are traversed in preorder.

b) The *postorder traversal* of T traverses in postorder the vertices of the subtrees T_1, $T_2, \ldots, T_k$ and then visits the root.

We demonstrate these ideas in the following example.

EXAMPLE 12.6

Consider the rooted tree shown in Fig. 12.19.

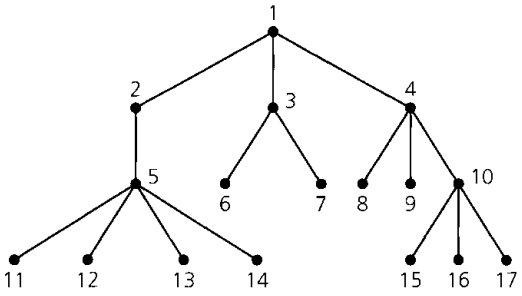


Figure 12.19

a) *Preorder*: After visiting vertex 1 we visit the subtree T_1 rooted at vertex 2. After visiting vertex 2 we proceed to the subtree rooted at vertex 5, and after visiting vertex 5 we go to the subtree rooted at vertex 11. This subtree has no other vertices, so we visit vertex 11 and then return to vertex 5 from which we visit, in succession, vertices 12, 13, and 14. Following this we *backtrack* (14 to 5 to 2 to 1) to the root and then visit the vertices in the subtree T_2 in the preorder 3, 6, 7. Finally, after returning to the root for the last time, we traverse the subtree T_3 in the preorder 4, 8, 9, 10, 15, 16, 17. Hence the preorder listing of the vertices in this tree is 1, 2, 5, 11, 12, 13, 14, 3, 6, 7, 4, 8, 9, 10, 15, 16, 17.

In this ordering we start at the root and build a path as far as we can. At each level we go to the leftmost vertex (not previously visited) at the next level, until we reach a leaf ℓ. Then we backtrack to the parent p of this leaf ℓ and visit ℓ's sibling s (and the subtree that s determines) directly on its right. If no such sibling s exists, we backtrack to the grandparent g of the leaf ℓ and visit, if it exists, a vertex u that is the sibling of p directly to its right in the tree. Continuing in this manner, we eventually visit (the first time each one is encountered) all of the vertices in the tree.

The vertices in Figs. 12.11(a), 12.12, and 12.17 are visited in preorder. The preorder traversal for the tree in Fig. 12.11(a) provides the ordering in Fig. 12.11(b). The lexicographic order in Example 12.4 arises from the preorder traversal of the tree in Fig. 12.12.

b) *Postorder*: For the postorder traversal of a tree, we start at the root r and build the longest path, going to the leftmost child of each internal vertex whenever we can. When we arrive at a leaf ℓ we visit this vertex and then backtrack to its parent p. However, we do not visit p until after all of its descendants are visited. The next vertex we visit is found by applying the same procedure at p that was originally applied at r in obtaining ℓ—except that now we first go from p to the sibling of ℓ directly to the right (of ℓ). And at no time is any vertex visited more than once or before any of its descendants.

For the tree given in Fig. 12.19, the postorder traversal starts with a postorder traversal of the subtree T_1 rooted at vertex 2. This yields the listing 11, 12, 13, 14, 5, 2. We proceed to the subtree T_2, and the postorder listing continues with 6, 7, 3. Then for T_3 we find 8, 9, 15, 16, 17, 10, 4 as the postorder listing. Finally, vertex 1 is visited. Consequently, for this tree, the postorder traversal visits the vertices in the order 11, 12, 13, 14, 5, 2, 6, 7, 3, 8, 9, 15, 16, 17, 10, 4, 1.

In the case of binary rooted trees, a third type of tree traversal called the inorder traversal may be used. Here we do *not* consider subtrees as first and second, but rather in terms of left and right. The formal definition is recursive, as were the definitions of preorder and postorder traversals.

Definition 12.4

Let $T = (V, E)$ be a binary rooted tree with vertex r the root.

1) If $|V| = 1$, then the vertex r constitutes the *inorder traversal* of T.
2) When $|V| > 1$, let T_L and T_R denote the left and right subtrees of T. The *inorder traversal* of T first traverses the vertices T_L in inorder, then it visits the root r, and then it traverses, in inorder, the vertices of T_R.

We realize that here a left or right subtree may be empty. Also, if v is a vertex in such a tree and $od(v) = 1$, then if w is the child of v, we must distinguish between w's being the left child and its being the right child.

EXAMPLE 12.7

As a result of the previous comments, the two binary rooted trees shown in Fig. 12.20 are not considered the same, when viewed as *ordered trees*. As rooted binary trees they are the same. (Each tree has the same set of vertices and the same set of directed edges.) However, when we consider the additional concept of left and right children, we see that in part (a) of the figure vertex v has right child a, whereas in part (b) vertex a is the left child of v. Consequently, when the difference between left and right children is taken into consideration, these trees are no longer viewed as the same tree.

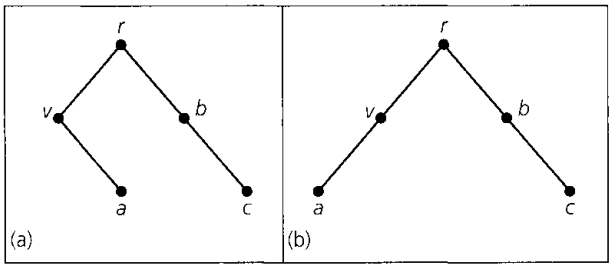


Figure 12.20

In visiting the vertices for the tree in part (a) of Fig. 12.20, we first visit in inorder the left subtree of the root r. This subtree consists of the root v and its *right child* a. (Here the left child is *null*, or nonexistent.) Since v has no left subtree, we visit in inorder vertex v and then its right subtree, namely, a. Having traversed the left subtree of r, we now visit vertex r and then traverse, in inorder, the vertices in the right subtree of r. This results in our visiting first vertex b (because b has no left subtree) and then vertex c. Hence the inorder listing for the tree shown in Fig. 12.20(a) is v, a, r, b, c.

When we consider the tree in part (b) of the figure, once again we start by visiting, in inorder, the vertices in the left subtree of the root r. Here, however, this left subtree consists of vertex v (the root of the subtree) and its *left* child a. (In this case, the right child of v is null, or nonexistent.) Therefore this inorder traversal first visits vertex a (the left subtree of v), and then vertex v. Since v has no right subtree, we are now finished visiting the left subtree of r, in inorder. So next the root r is visited, and then the vertices of the right subtree

of r are traversed, in inorder. This results in the inorder listing a, v, r, b, c for the tree shown in Fig. 12.20(b).

We should note, however, that for the preorder traversal *in this particular example*,[†] the same result is obtained for both trees:

$$\text{Preorder listing:} \quad r, v, a, b, c.$$

Likewise, this particular example is such that the postorder traversal for either tree gives us the following:

$$\text{Postorder listing:} \quad a, v, c, b, r.$$

It is only for the inorder traversal, with its distinctions between left and right children and between left and right subtrees, that a difference occurs. For the trees in parts (a) and (b) of Fig. 12.20 we found the respective inorder listings to be

$$\text{(a) } v, a, r, b, c \qquad \text{and} \qquad \text{(b) } a, v, r, b, c.$$

EXAMPLE 12.8

If we apply the inorder traversal to the binary rooted tree shown in Fig. 12.21, we find that the inorder listing for the vertices is $p, j, q, f, c, k, g, a, d, r, b, h, s, m, e, i, t, n, u$.

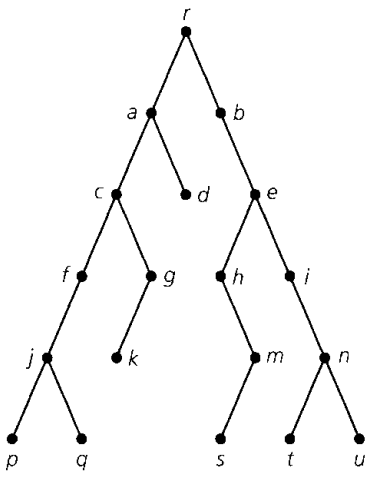


Figure 12.21

Our next example shows how the preorder traversal can be used in a counting problem dealing with binary trees.

EXAMPLE 12.9[‡]

For $n \geq 0$, consider the complete binary trees on $2n + 1$ vertices. The cases for $0 \leq n \leq 3$ are shown in Fig. 12.22. Here we distinguish left from right. So, for example, the two

---

[†]A note of caution! If we interchange the order of the two existing children (of a certain parent) in a binary rooted tree, then a change results in the preorder, postorder, and inorder traversals. If one child is "null," however, then only the inorder traversal changes.

[‡]This example uses material developed in the optional Sections 1.5 and 10.5. It may be omitted with no loss of continuity.

complete binary trees for $n = 2$ are considered distinct. [If we do not distinguish left from right, these trees are (isomorphic and) no longer counted as two different trees.]

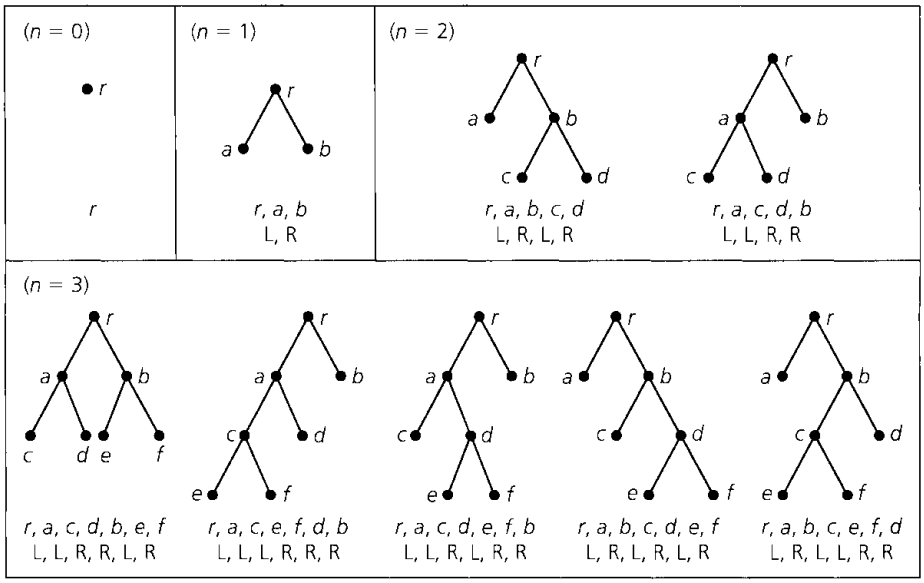


Figure 12.22

Below each tree in the figure we list the vertices for a preorder traversal. In addition, for $1 \leq n \leq 3$, we find a list of n L's and n R's under each preorder traversal. These lists are determined as follows. The first tree for $n = 2$, for instance, has the list L, R, L, R because, after visiting the root r, we go to the left (L) subtree rooted at a and visit vertex a. Then we backtrack to r and go to the right (R) subtree rooted at b. After visiting vertex b we go to the left (L) subtree of b rooted at c and visit vertex c. Then, lastly, we backtrack to b and go to its right (R) subtree to visit vertex d. This generates the list L, R, L, R and the other seven lists of L's and R's are obtained in the same way.

Since we are traversing these trees in preorder, each list starts with an L. There is an equal number of L's and R's in each list because the trees are complete binary trees. Finally, the number of R's never exceeds the number of L's as a given list is read from left to right — again, because we have a preorder traversal. Should we replace each L by a 1 and each R by a -1, for the five trees for $n = 3$, we find ourselves back in part (a) of Example 1.43, where we have one of our early examples of the Catalan numbers. Hence, for $n \geq 0$, we see that the number of complete binary trees on $2n + 1$ vertices is $\frac{1}{n+1}\binom{2n}{n}$, the nth Catalan number. [Note that if we *prune* the five trees for $n = 3$ by removing the four leaves for each tree, we obtain the five rooted ordered binary trees in Fig. 10.18.]

The notion of preorder now arises in the following procedure for finding a spanning tree for a connected graph.

Let $G = (V, E)$ be a loop-free connected undirected graph with $r \in V$. Starting from r, we construct a path in G that is as long as possible. If this path includes every vertex in V, then the path is a spanning tree T for G and we are finished. If not, let x and y be the last two vertices visited along this path, with y the last vertex. We then return, or *backtrack*, to the vertex x and construct a second path in G that is as long as possible, starts at x, and

doesn't include any vertex already visited. If no such path exists, backtrack to the parent p of x and see how far it is possible to branch off from p, building a path (that is as long as possible and has no previously visited vertices) to a new vertex y_1 (which will be a new leaf for T). Should all edges from the vertex p lead to vertices already encountered, backtrack one level higher and continue the process. Since the graph is finite and connected, this technique, which is called *backtracking*, or *depth-first search*, eventually determines a spanning tree T for G, where r is regarded as the root of T. Using T, we then order the vertices of G in a preorder listing.

The depth-first search serves as a framework around which many algorithms can be designed to test for certain graph properties. One such algorithm will be examined in detail in Section 12.5.

One way to help implement the depth-first search in a computer program is to assign a fixed order to the vertices of the given graph $G = (V, E)$. Then if there are two or more vertices adjacent to a vertex v and none of these vertices has already been visited, we shall know exactly which vertex to visit first. This order now helps us to develop the foregoing description of the depth-first search as an algorithm.

Let $G = (V, E)$ be a loop-free connected undirected graph where $|V| = n$ and the vertices are ordered as $v_1, v_2, v_3, \ldots, v_n$. To find the rooted ordered depth-first spanning tree for the prescribed order, we apply the following algorithm, wherein the variable v is used to store the vertex presently being examined.

Depth-First Search Algorithm

Step 1: Assign v_1 to the variable v and initialize T as the tree consisting of just this one vertex. (The vertex v_1 will be the root of the spanning tree that develops.) Visit v_1.

Step 2: Select the smallest subscript i, for $2 \leq i \leq n$, such that $\{v, v_i\} \in E$ and v_i has not already been visited.

If no such subscript is found, then go to step (3). Otherwise, perform the following: (1) Attach the edge $\{v, v_i\}$ to the tree T and visit v_i; (2) Assign v_i to v; and (3) Return to step (2).

Step 3: If $v = v_1$, the tree T is the (rooted ordered) spanning tree for the order specified.

Step 4: For $v \neq v_1$, backtrack from v to its parent u in T. Then assign u to v and return to step (2).

EXAMPLE 12.10

We now apply this algorithm to the graph $G = (V, E)$ shown in Fig. 12.23(a). Here the order for the vertices is alphabetic: $a, b, c, d, e, f, g, h, i, j$.

First we assign the vertex a to the variable v and initialize T as just the vertex a (the root). We visit vertex a. Then, going to step (2), we find that the vertex b is the first vertex w such that $\{a, w\} \in E$ and w has not been visited earlier. So we attach edge $\{a, b\}$ to T and visit b, assign b to v, and then return to step (2).

At $v = b$ we find that the first vertex (not visited earlier) that provides an edge for the spanning tree is d. Consequently, the edge $\{b, d\}$ is attached to T and d is visited, then d is assigned to v, and we again return to step (2).

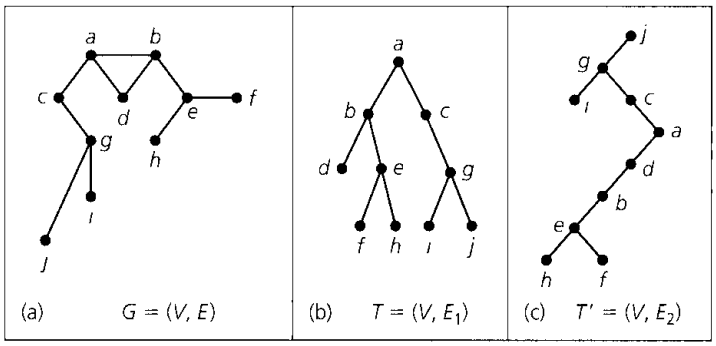


Figure 12.23

This time, however, there is no new vertex that we can obtain from d, because vertices a and b have already been visited. So we go to step (3). But here the value of v is d, not a, and we go to step (4). Now we backtrack from d, assigning the vertex b to v, and then we return to step (2). At this time we add the edge $\{b, e\}$ to T and visit e.

Continuing the process, we attach the edge $\{e, f\}$ (and visit f) and then the edge $\{e, h\}$ (and visit h). But now the vertex h has been assigned to v, and we must backtrack from h to e to b to a. When v is assigned the vertex a this (second) time, the new edge $\{a, c\}$ is obtained and vertex c is visited. Then we proceed to attach the edges $\{c, g\}$, $\{g, i\}$, and $\{g, j\}$ (visiting the vertices g, i, and j, respectively). At this point all of the vertices in G have been visited, and we backtrack from j to g to c to a. With $v = a$ once again we return to step (2) and from there to step (3), where the process terminates.

The resulting tree $T = (V, E_1)$ is shown in part (b) of Fig. 12.23. Part (c) of the figure shows the tree T' that results for the vertex ordering: $j, i, h, g, f, e, d, c, b, a$.

A second method for searching the vertices of a loop-free connected undirected graph is the *breadth-first search*. Here we designate one vertex as the root and fan out to all vertices adjacent to the root. From each child of the root we then fan out to those vertices (not previously visited) that are adjacent to one of these children. As we continue this process, we never list a vertex twice, so no cycle is constructed, and with G finite the process eventually terminates.

We actually used this technique earlier in Example 11.28 of Section 11.5.

A certain data structure proves useful in developing an algorithm for this second searching method. A *queue* is an ordered list wherein items are inserted at one end (called the *rear*) of the list and deleted at the other end (called the *front*). The *first* item inserted *in* the queue is the *first* item that can be taken out of it. Consequently, a queue is referred to as a "first-in, first-out," or FIFO, structure.

As in the depth-first search, we again assign an order to the vertices of our graph.

We start with a loop-free connected undirected graph $G = (V, E)$, where $|V| = n$ and the vertices are ordered as $v_1, v_2, v_3, \ldots, v_n$. The following algorithm generates the (rooted ordered) breadth-first spanning tree T of G for the given order.

Breadth-First Search Algorithm

Step 1: Insert vertex v_1 at the rear of the (initially empty) queue Q and initialize T as the tree made up of this one vertex v_1 (the root of the final version of T). Visit v_1.

Step 2: While the queue Q is not empty, delete the vertex v from the front of Q. Now examine the vertices v_i (for $2 \leq i \leq n$) that are adjacent to v — in the specified order. If v_i has not been visited, perform the following: (1) Insert v_i at the rear of Q; (2) Attach the edge $\{v, v_i\}$ to T; and (3) Visit vertex v_i. [If we examine *all* of the vertices previously in the queue Q and obtain no new edges, then the tree T (generated to this point) is the (rooted ordered) spanning tree for the given order.]

EXAMPLE 12.11

We shall employ the graph of Fig. 12.23(a) with the prescribed order $a, b, c, d, e, f, g, h, i, j$ to illustrate the use of the algorithm for the breadth-first search.

Start with vertex a. Insert a at the rear of (the presently empty) queue Q, initialize T as this one vertex (the root of the resulting tree), and visit vertex a.

In step (2) we now delete a from (the front of) Q and examine the vertices adjacent to a — namely, the vertices b, c, d. (These vertices have not been previously visited.) This results in our (i) inserting vertex b at the rear of Q, attaching the edge $\{a, b\}$ to T, and visiting vertex b; (ii) inserting vertex c at the rear of Q (after b), attaching the edge $\{a, c\}$ to T, and visiting vertex c; and (iii) inserting vertex d at the rear of Q (after c), attaching the edge $\{a, d\}$ to T, and visiting vertex d.

Since the queue Q is not empty, we execute step (2) again. Upon deleting vertex b from the front of Q, we now find that the only vertex adjacent to b (that has not been previously visited) is e. So we insert vertex e at the rear of Q (after d), attach the edge $\{b, e\}$ to T, and visit vertex e. Continuing with vertex c we obtain the new (unvisited) vertex g. So we insert vertex g at the rear of Q (after e), attach the edge $\{c, g\}$ to T, and visit vertex g. And now we delete vertex d from the front of Q. But at this point there are no unvisited vertices adjacent to d, so we then delete vertex e from the front of Q. This vertex leads to the following: inserting vertex f at the rear of Q (after g), attaching the edge $\{e, f\}$ to T, and visiting vertex f. This is followed by: inserting vertex h at the rear of Q (after f), attaching edge $\{e, h\}$ to T, and visiting vertex h. Continuing with vertex g, we insert vertex i at the rear of Q (after h), attach edge $\{g, i\}$ to T, and visit vertex i, and then we insert vertex j at the rear of Q (after i), attach edge $\{g, j\}$ to T, and visit vertex j.

Once again we return to the beginning of step (2). But now when we delete (from the front of Q) and examine each of the vertices f, h, i, and j (in this order), we find no unvisited vertices for any of these four vertices. Consequently, the queue Q now remains empty and the tree T in Fig. 12.24(a) is the breadth-first spanning tree for G, for the order

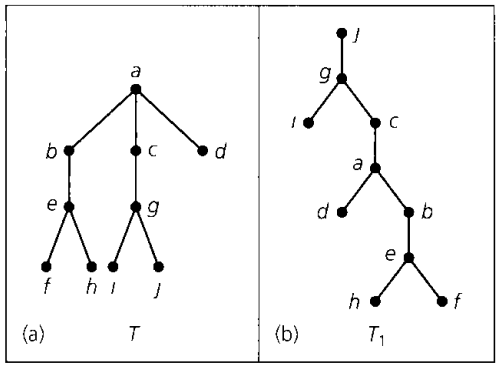


Figure 12.24

prescribed. (The tree T_1, shown in part (b) of the figure, arises for the order $j, i, h, g, f, e, d, c, b, a$.)

Let us apply these ideas on graph searching to one more example.

EXAMPLE 12.12

Let $G = (V, E)$ be an undirected graph (with loops) where the vertices are ordered as $v_1, v_2, \ldots, v_7$. If Fig. 12.25(a) is the adjacency matrix $A(G)$ for G, how can we use this representation of G to determine whether G is connected, without drawing the graph?

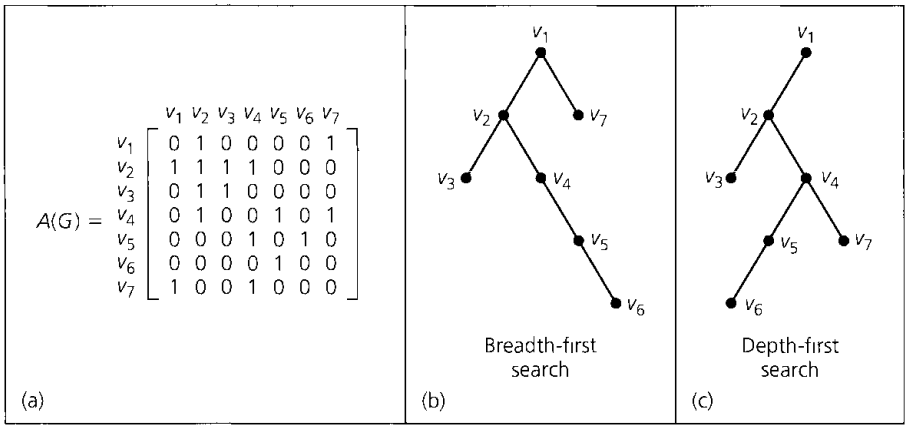


Figure 12.25

Using v_1 as the root, in part (b) of the figure we search the graph by means of its adjacency matrix, using a breadth-first search. [Here we ignore the loops by ignoring any 1's on the main diagonal (extending from the upper left to the lower right).] First we visit the vertices adjacent to v_1, listing them in ascending order according to the subscripts on the v's in $A(G)$. The search continues, and as all vertices in G are reached, G is shown to be connected.

The same conclusion follows from the depth-first search in part (c). The tree here also has v_1 as its root. As the tree branches out to search the graph, it does so by listing the first vertex found adjacent to v_1 according to the row in $A(G)$ for v_1. Likewise, from v_2 the first new vertex in this search is found from $A(G)$ to be v_3. The vertex v_3 is a leaf in this tree because no new vertex can be visited from v_3. As we backtrack to v_2, row 2 of $A(G)$ indicates that v_4 can now be visited from v_2. As this process continues, the connectedness of G follows from part (c) of the figure.

It is time now to return to our main discussion on rooted trees. The following definition generalizes the ideas that were introduced for Example 12.5.

Definition 12.5

Let $T = (V, E)$ be a rooted tree, and let $m \in \mathbf{Z}^+$.

We call T an *m-ary tree* if $od(v) \leq m$ for all $v \in V$. When $m = 2$, the tree is called a *binary tree*.

If $od(v) = 0$ or m, for all $v \in V$, then T is called a *complete m-ary tree*. The special case of $m = 2$ results in a *complete binary tree*.

In a complete m-ary tree, each internal vertex has exactly m children. (Each leaf of this tree still has no children.)

Some properties of these trees are considered in the following theorem.

THEOREM 12.6

Let $T = (V, E)$ be a complete m-ary tree with $|V| = n$. If T has ℓ leaves and i internal vertices, then (a) $n = mi + 1$; (b) $\ell = (m - 1)i + 1$; and (c) $i = (\ell - 1)/(m - 1) = (n - 1)/m$.

Proof: This proof is left for the Section Exercises.

EXAMPLE 12.13

The Wimbledon tennis championship is a single-elimination tournament wherein a player (or doubles team) is eliminated after a single loss. If 27 women compete in the singles championship, how many matches must be played to determine the number-one female player?

Consider the tree shown in Fig. 12.26. With 27 women competing, there are 27 leaves in this complete binary tree, so from Theorem 12.6(c) the number of internal vertices (which is the number of matches) is $i = (\ell - 1)/(m - 1) = (27 - 1)/(2 - 1) = 26$.

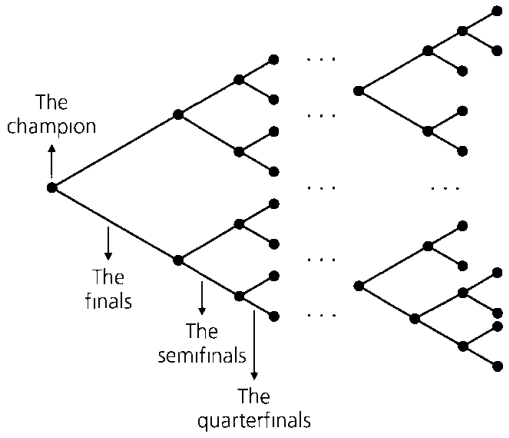


Figure 12.26

EXAMPLE 12.14

A classroom contains 25 microcomputers that must be connected to a wall socket that has four outlets. Connections are made by using extension cords that have four outlets each. What is the least number of cords needed to get these computers set up for class use?

The wall socket is considered the root of a complete m-ary tree for $m = 4$. The microcomputers are the leaves of this tree, so $\ell = 25$. Each internal vertex, except the root, corresponds with an extension cord. So by part (c) of Theorem 12.6, there are $(\ell - 1)/(m - 1) = (25 - 1)/(4 - 1) = 8$ internal vertices. Hence we need $8 - 1$ (where the 1 is subtracted for the root) $= 7$ extension cords.

Definition 12.6

If $T = (V, E)$ is a rooted tree and h is the largest level number achieved by a leaf of T, then T is said to have *height* h. A rooted tree T of height h is said to be *balanced* if the level number of every leaf in T is $h - 1$ or h.

The rooted tree shown in Fig. 12.19 is a balanced tree of height 3. Tree T' in Fig. 12.23(c) has height 7 but is not balanced. (Why?)

The tree for the tournament in Example 12.13 must be balanced so that the tournament will be as fair as possible. If it is not balanced, some competitor will receive more than one bye (an opportunity to advance without playing a match).

Before stating our next theorem, let us recall that for all $x \in \mathbf{R}$, $\lfloor x \rfloor$ denotes the greatest integer in x, or floor of x, whereas $\lceil x \rceil$ designates the ceiling of x.

THEOREM 12.7

Let $T = (V, E)$ be a complete m-ary tree of height h with ℓ leaves. Then $\ell \leq m^h$ and $h \geq \lceil \log_m \ell \rceil$.

Proof: The proof that $\ell \leq m^h$ will be established by induction on h. When $h = 1$, T is a tree with a root and m children. In this case $\ell = m = m^h$, and the result is true. Assume the result true for all trees of height $< h$, and consider a tree T with height h and ℓ leaves. (The level numbers that are possible for these leaves are $1, 2, \ldots, h$, with at least m of the leaves at level h.) The ℓ leaves of T are also the ℓ leaves (total) for the m subtrees T_i, $1 \leq i \leq m$, of T rooted at each of the children of the root. For $1 \leq i \leq m$, let ℓ_i be the number of leaves in subtree T_i. (In the case where leaf and root coincide, $\ell_i = 1$. But since $m \geq 1$ and $h - 1 \geq 0$, we have $m^{h-1} \geq 1 = \ell_i$.) By the induction hypothesis, $\ell_i \leq m^{h(T_i)} \leq m^{h-1}$, where $h(T_i)$ denotes the height of the subtree T_i, and so $\ell = \ell_1 + \ell_2 + \cdots + \ell_m \leq m(m^{h-1}) = m^h$.

With $\ell \leq m^h$, we find that $\log_m \ell \leq \log_m(m^h) = h$, and since $h \in \mathbf{Z}^+$, it follows that $h \geq \lceil \log_m \ell \rceil$.

COROLLARY 12.1

Let T be a balanced complete m-ary tree with ℓ leaves. Then the height of T is $\lceil \log_m \ell \rceil$.

Proof: This proof is left as an exercise.

We close this section with an application that uses a complete ternary ($m = 3$) tree.

EXAMPLE 12.15

Decision Trees. There are eight coins (identical in appearance) and a pan balance. If exactly one of these coins is counterfeit and heavier than the other seven, find the counterfeit coin.

Let the coins be labeled $1, 2, 3, \ldots, 8$. In using the pan balance to compare sets of coins there are three outcomes to consider: (a) the two sides balance to indicate that the coins in the two pans are not counterfeit; (b) the left pan of the balance goes down, indicating that the counterfeit coin is in the left pan; or (c) the right pan goes down, indicating that it holds the counterfeit coin.

In Fig. 12.27(a), we search for the counterfeit coin by first balancing coins 1, 2, 3, 4 against 5, 6, 7, 8. If the balance tips to the right, we follow the right branch from the root to then analyze coins 5, 6 against 7, 8. If the balance tips to the left, we test coins 1, 2 against 3, 4. At each successive level, we have half as many coins to test, so at level 3 (after three weighings) the heavier counterfeit coin has been identified.

The tree in part (b) of the figure finds the heavier coin in two weighings. The first weighing balances coins 1, 2, 3 against 6, 7, 8. Three possible outcomes can occur: (i) the balance tips to the right, indicating that the heavier coin is 6, 7, or 8, and we follow the right branch from the root; (ii) the balance tips to the left and we follow the left branch to find which of 1, 2, 3 is the heavier; or (iii) the pans balance and we follow the center branch to find which of 4, 5 is heavier. At each internal vertex the label indicates which coins are being compared.

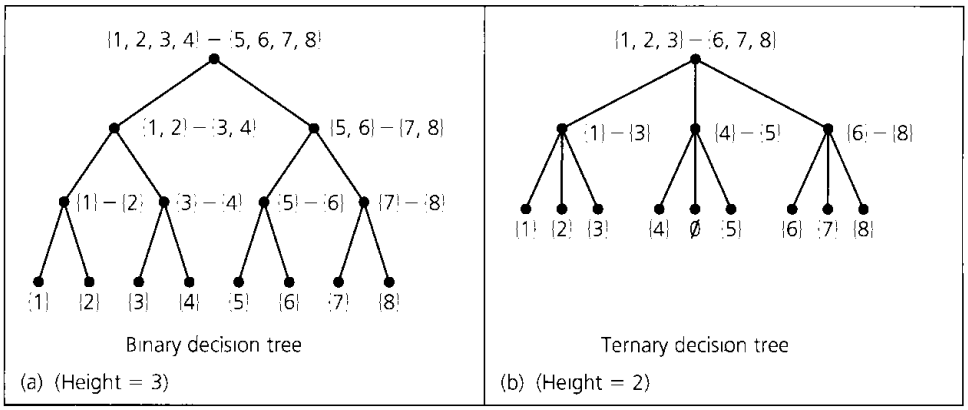


Figure 12.27

Unlike part (a), a conclusion may be deduced in part (b) when a coin is not included in a weighing. Finally, when comparing coins 4 and 5, because equality cannot take place we label the center leaf with $\emptyset$.

In this particular problem, we claim that the height of the complete ternary tree used must be at least 2. With eight coins involved, the tree will have at least eight leaves. Consequently, with $\ell \geq 8$, it follows from Theorem 12.7 that $h \geq \lceil \log_3 \ell \rceil \geq \lceil \log_3 8 \rceil = 2$, so at least two weighing are needed. If n coins are involved, the complete ternary tree will have ℓ leaves where $\ell \geq n$, and its height h satisfies $h \geq \lceil \log_3 n \rceil$.

EXERCISES 12.2

1. Answer the following questions for the tree shown in Fig. 12.28.

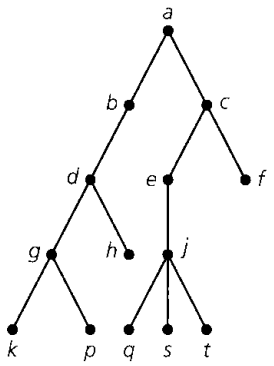


Figure 12.28

a) Which vertices are the leaves?

b) Which vertex is the root?

c) Which vertex is the parent of g?

d) Which vertices are the descendants of c?

e) Which vertices are the siblings of s?

f) What is the level number of vertex f?

g) Which vertices have level number 4?

2. Let $T = (V, E)$ be a binary tree. In Fig. 12.29 we find the subtree of T rooted at vertex p. (The dashed line coming into vertex p indicates that there is more to the tree T than what appears in the figure.) If the level number for vertex u is 37, (a) what are the level numbers for vertices p, s, t, v, w, x, y, and z? (b) how many ancestors does vertex u have? (c) how many ancestors does vertex y have?

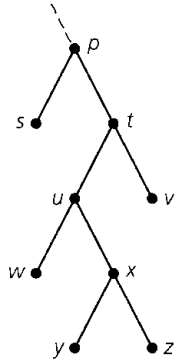


Figure 12.29

3. a) Write the expression $(w + x - y)/(\pi * z^3)$ in Polish notation, using a rooted tree.

b) What is the value of the expression (in Polish notation) $/ \uparrow a - bc + d * ef$, if $a = c = d = e = 2, b = f = 4$?

4. Let $T = (V, E)$ be a rooted tree ordered by a universal address system. (a) If vertex v in T has address 2.1.3.6, what is the smallest number of siblings that v must have? (b) For the vertex v in part (a), find the address of its parent. (c) How many ancestors does the vertex v in part (a) have? (d) With the presence of v in T, what other addresses must there be in the system?

5. For the tree shown in Fig. 12.30, list the vertices according to a preorder traversal, an inorder traversal, and a postorder traversal.

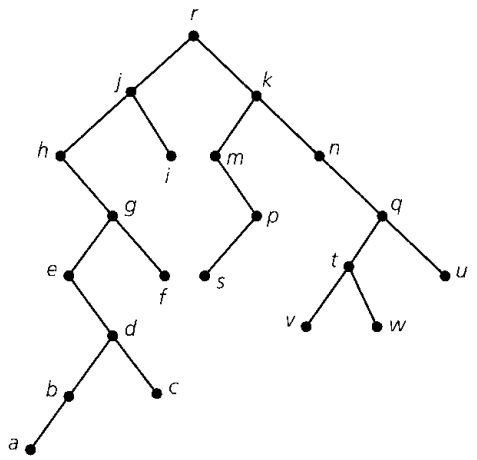


Figure 12.30

6. List the vertices in the tree shown in Fig. 12.31 when they are visited in a preorder traversal and in a postorder traversal.

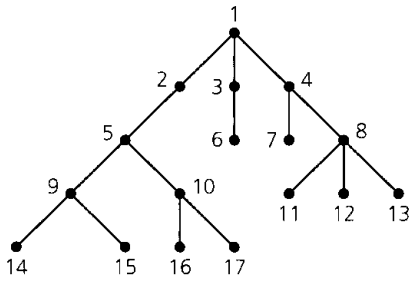


Figure 12.31

7. a) Find the depth-first spanning tree for the graph shown in Fig. 11.72(a) if the order of the vertices is given as (i) a, b, c, d, e, f, g, h; (ii) h, g, f, e, d, c, b, a; (iii) a, b, c, d, h, g, f, e.

b) Repeat part (a) for the graph shown in Fig. 11.85(i).

8. Find the breadth-first spanning trees for the graphs and prescribed orders given in Exercise 7.

9. Let $G = (V, E)$ be an undirected graph with adjacency matrix $A(G)$ as shown here.

$$A(G) = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \end{array} \begin{array}{c} \begin{array}{cccccccc} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & v_7 & v_8 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Use a breadth-first search based on $A(G)$ to determine whether G is connected.

10. a) Let $T = (V, E)$ be a binary tree. If $|V| = n$, what is the maximum height that T can attain?

b) If $T = (V, E)$ is a complete binary tree and $|V| = n$, what is the maximum height that T can reach in this case?

11. Prove Theorem 12.6 and Corollary 12.1.

12. With m, n, i, ℓ as in Theorem 12.6, prove that

a) $n = (m\ell - 1)/(m - 1)$. **b)** $\ell = [(m - 1)n + 1]/m$.

13. a) A complete ternary (or 3-ary) tree $T = (V, E)$ has 34 internal vertices. How many edges does T have? How many leaves?

b) How many internal vertices does a complete 5-ary tree with 817 leaves have?

14. The complete binary tree $T = (V, E)$ has $V = \{a, b, c, \ldots, i, j, k\}$. The postorder listing of V yields $d, e, b, h, i, f, j, k, g, c, a$. From this information draw T if (a) the height of T is 3; (b) the height of the left subtree of T is 3.

15. For $m \geq 3$, a complete m-ary tree can be transformed into a complete binary tree by applying the idea shown in Fig. 12.32.

a) Use this technique to transform the complete ternary decision tree shown in Fig. 12.27(b).

b) If T is a complete quaternary tree of height 3, what is the maximum height that T can have after it is transformed into a complete binary tree? What is the minimum height?

c) Answer part (b) if T is a complete m-ary tree of height h.

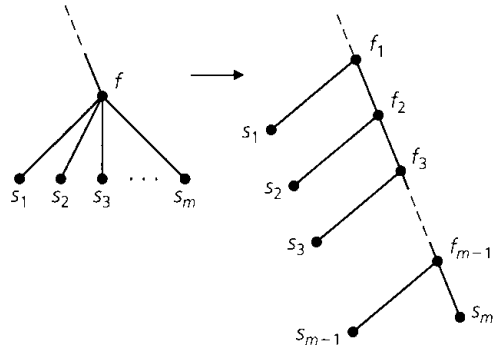


Figure 12.32

16. **a)** At a men's singles tennis tournament, each of 25 players brings a can of tennis balls. When a match is played, one can of balls is opened and used, then kept by the loser. The winner takes the unopened can on to his next match. How many cans of tennis balls will be opened during this tournament? How many matches are played in the tournament?

b) In how many matches did the tournament champion play?

17. What is the maximum number of internal vertices that a complete quaternary tree of height 8 can have? What is the number for a complete m-ary tree of height h?

18. On the first Sunday of 2003 Rizzo and Frenchie start a chain letter, each of them sending five letters (to ten different friends between them). Each person receiving the letter is to send five copies to five new people on the Sunday following the letter's arrival. After the first seven Sundays have passed, what is the total number of chain letters that have been mailed? How many were mailed on the last three Sundays?

19. Use a complete ternary decision tree to repeat Example 12.15 for a set of 12 coins, exactly one of which is heavier (and counterfeit).

20. Let $T = (V, E)$ be a balanced complete m-ary tree of height $h \geq 2$. If T has ℓ leaves and b_{h-1} internal vertices at level $h - 1$, explain why $\ell = m^{h-1} + (m - 1)b_{h-1}$.

21. Consider the complete binary trees on 31 vertices. (Here we distinguish left from right as in Example 12.9.) How many of these trees have 11 vertices in the left subtree of the root? How many have 21 vertices in the right subtree of the root?

22. For $n \geq 0$, let a_n count the number of complete binary trees on $2n + 1$ vertices. (Here we distinguish left from right as in Example 12.9.) How is a_{n+1} related to $a_0, a_1, a_2, \ldots, a_{n-1}, a_n$?

23. Consider the following algorithm where the input is a rooted tree with root r.

Step 1: Push r onto the (empty) stack
Step 2: While the stack is not empty
  Pop the vertex at the top of
    the stack and record its label
  Push the children — going from
    right to left — of this vertex
    onto the stack

(The stack data structure was explained in Example 10.43).

What is the output when this algorithm is applied to (a) the tree in Fig. 12.19? (b) any rooted tree?

24. Consider the following algorithm where the input is a rooted tree with root r.

Step 1: Push r onto the (empty) stack
Step 2: While the stack is not empty
  If the entry at the top of the stack is
    not marked
    Then mark it and push its
      children — right to left — onto
      the stack
  Else
    Pop the vertex at the top of the
      stack and record its label

What is the output when the algorithm is applied to (a) the tree in Fig. 12.19? (b) any rooted tree?

12.3 Trees and Sorting

In Example 10.5, the bubble sort was introduced. There we found that the number of comparisons needed to sort a list of n items is $n(n - 1)/2$. Consequently, this algorithm determines a function $h: \mathbf{Z}^+ \to \mathbf{R}$ defined by $h(n) = n(n - 1)/2$. This is the (worst-case) time-complexity function for the algorithm, and we often express this by writing $h \in O(n^2)$. Consequently, the bubble sort is said to require $O(n^2)$ comparisons. We interpret this to mean that for large n, the number of comparisons is bounded above by cn^2, where c is a constant that is generally not specified because it depends on such factors as the compiler and the computer that are used.

In this section we shall study a second method for sorting a given list of n items into ascending order. The method is called the *merge sort*, and we shall find that the order of its worst-case time-complexity function is $O(n \log_2 n)$. This will be accomplished in the following manner:

1) First we shall measure the number of comparisons needed when n is a power of 2. Our method will employ a pair of balanced complete binary trees.

2) Then we shall cover the case for general n by using the optional material on divide-and-conquer algorithms in Section 10.6.

For the case where n is an arbitrary positive integer, we start by considering the following procedure.

Given a list of n items to sort into ascending order, the *merge sort* recursively splits the given list and all subsequent sublists in half (or as close as possible to half) until each sublist contains a single element. Then the procedure merges these sublists in ascending order until the original n items have been so sorted. The splitting and merging processes can best be described by a pair of balanced complete binary trees, as in the next example.

EXAMPLE 12.16

Merge Sort. Using the merge sort, Fig. 12.33 sorts the list 6, 2, 7, 3, 4, 9, 5, 1, 8. The tree at the top of the figure shows how the process first splits the given list into sublists of size 1. The merging process is then outlined by the tree at the bottom of the figure.

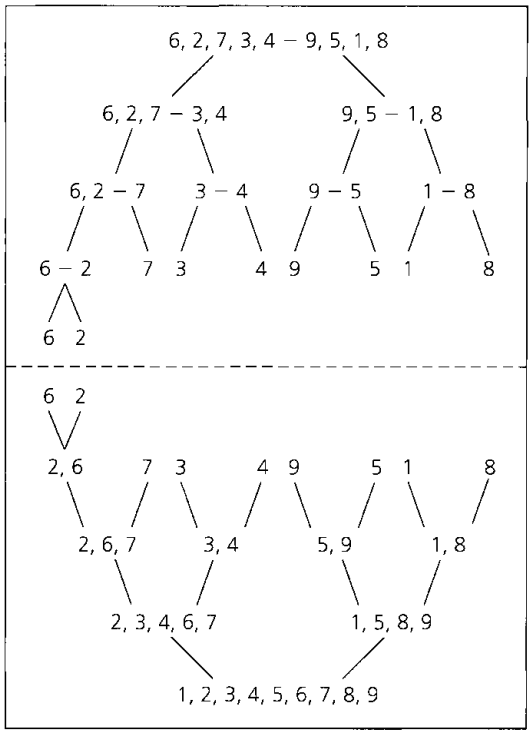


Figure 12.33

To compare the merge sort to the bubble sort, we want to determine its (worst-case) time-complexity function. The following lemma will be needed for this task.

LEMMA 12.1

Let L_1 and L_2 be two sorted lists of ascending numbers, where L_i contains n_i elements, for $i = 1, 2$. Then L_1 and L_2 can be merged into one ascending list L using at most $n_1 + n_2 - 1$ comparisons.

Proof: To merge L_1, L_2 into list L, we perform the following algorithm.

Step 1: Set L equal to the empty list $\emptyset$.

Step 2: Compare the first elements in L_1, L_2. Remove the smaller of the two from the list it is in and place it at the end of L.

Step 3: For the *present* lists L_1, L_2 [one change is made in one of these lists each time step (2) is executed], there are two considerations.

a) If either of L_1, L_2 is empty, then the other list is concatenated to the end of L. This completes the merging process.

b) If not, return to step (2).

Each comparison of a number from L_1 with one from L_2 results in the placement of an element at the end of list L, so there cannot be more than $n_1 + n_2$ comparisons. When one of the lists L_1 or L_2 becomes empty no further comparisons are needed, so the maximum number of comparisons needed is $n_1 + n_2 - 1$.

To determine the (worst-case) time-complexity function of the merge sort, consider a list of n elements. For the moment, we do not treat the general problem, assuming here that $n = 2^h$.[†] In the splitting process, the list of 2^h elements is first split into two sublists of size 2^{h-1}. (These are the level 1 vertices in the tree representing the splitting process.) As the process continues, each successive list of size 2^{h-k}, $h > k$, is at level k and splits into two sublists of size $(1/2)(2^{h-k}) = 2^{h-k-1}$. At level h the sublists each contain $2^{h-h} = 1$ element.

Reversing the process, we first merge the $n = 2^h$ leaves into 2^{h-1} ordered sublists of size 2. These sublists are at level $h - 1$ and require $(1/2)(2^h) = 2^{h-1}$ comparisons (one per pair). As this merging process continues, at each of the 2^k vertices at level k, $1 \leq k < h$, there is a sublist of size 2^{h-k}, obtained from merging the two sublists of size 2^{h-k-1} at its children (on level $k + 1$). From Lemma 12.1, this merging requires at most $2^{h-k-1} + 2^{h-k-1} - 1 = 2^{h-k} - 1$ comparisons. When the children of the root are reached, there are two sublists of size 2^{h-1} (at level 1). To merge these sublists into the final list requires at most $2^{h-1} + 2^{h-1} - 1 = 2^h - 1$ comparisons.

Consequently, for $1 \leq k \leq h$, at level k there are 2^{k-1} pairs of vertices. At each of these vertices is a sublist of size 2^{h-k}, so it takes at most $2^{h-k+1} - 1$ comparisons to merge each pair of sublists. With 2^{k-1} pairs of vertices at level k, the total number of comparisons at level k is at most $2^{k-1}(2^{h-k+1} - 1)$. When we sum over all levels k, where $1 \leq k \leq h$, we find that the total number of comparisons is at most

$$\sum_{k=1}^{h} 2^{k-1}(2^{h-k+1} - 1) = \sum_{k=0}^{h-1} 2^k(2^{h-k} - 1) = \sum_{k=0}^{h-1} 2^h - \sum_{k=0}^{h-1} 2^k = h \cdot 2^h - (2^h - 1).$$

With $n = 2^h$, we have $h = \log_2 n$ and

$$h \cdot 2^h - (2^h - 1) = n \log_2 n - (n - 1) = n \log_2 n - n + 1,$$

[†] The result obtained here for $n = 2^h$, $h \in \mathbf{N}$, is actually true for all $n \in \mathbf{Z}^+$. However, the derivation for general n requires the optional material in Section 10.6. That is why this counting argument is included here—for the benefit of those readers who did not cover Section 10.6.

where $n \log_2 n$ is the dominating term for large n. Thus the (worst-case) time-complexity function for this sorting procedure is $g(n) = n \log_2 n - n + 1$ and $g \in O(n \log_2 n)$, for $n = 2^h$, $h \in \mathbf{Z}^+$. Hence the number of comparisons needed to merge sort a list of n items is bounded above by $dn \log_2 n$ for some constant d, and for all $n \geq n_0$, where n_0 is some particular (large) positive integer.

To show that the order of the merge sort is $O(n \log_2 n)$ for all $n \in \mathbf{Z}^+$, our second approach will use the result of Exercise 9 from Section 10.6. We state that now:

Let $a, b, c \in \mathbf{Z}^+$, with $b \geq 2$. If $g: \mathbf{Z}^+ \to \mathbf{R}^+ \cup \{0\}$ is a monotone increasing function, where

$$g(1) \leq c,$$
$$g(n) \leq ag(n/b) + cn, \quad \text{for } n = b^h, h \in \mathbf{Z}^+,$$

then for the case where $a = b$, we have $g \in O(n \log n)$, for *all* $n \in \mathbf{Z}^+$. (The base for the log function may be any real number greater than 1. Here we shall use the base 2.)

Before we can apply this result to the merge sort, we wish to formulate this sorting process (illustrated in Fig. 12.33) as a precise algorithm. To do so, we call the procedure outlined in Lemma 12.1 the "merge" algorithm. Then we shall write "merge (L_1, L_2)" in order to represent the application of that procedure to the lists L_1, L_2, which are in ascending order.

The algorithm for merge sort is a recursive procedure because it may invoke itself. Here the input is an array (called List) of n items, such as real numbers.

The MergeSort Algorithm

Step 1: If $n = 1$, then List is already sorted and the process terminates. If $n > 1$, then go to step (2).

Step 2: (Divide the array and sort the subarrays.) Perform the following:

1) Assign m the value $\lfloor n/2 \rfloor$.

2) Assign to List 1 the subarray

$$\text{List}[1], \text{List}[2], \ldots, \text{List}[m].$$

3) Assign to List 2 the subarray

$$\text{List}[m + 1], \text{List}[m + 2], \ldots, \text{List}[n].$$

4) Apply MergeSort to List 1 (of size m) and to List 2 (of size $n - m$).

Step 3: Merge (List 1, List 2).

The function $g: \mathbf{Z}^+ \to \mathbf{R}^+ \cup \{0\}$ will measure the (worst-case) time-complexity for this algorithm by counting the maximum number of comparisons needed to merge sort an array of n items. For $n = 2^h$, $h \in \mathbf{Z}^+$, we have

$$g(n) = 2g(n/2) + [(n/2) + (n/2) - 1].$$

The term $2g(n/2)$ results from step (2) of the MergeSort algorithm, and the summand $[(n/2) + (n/2) - 1]$ follows from step (3) of the algorithm and Lemma 12.1.

With $g(1) = 0$, the preceding equation provides the inequalities

$$g(1) = 0 \leq 1,$$
$$g(n) = 2g(n/2) + (n-1) \leq 2g(n/2) + n, \quad \text{for } n = 2^h,\ h \in \mathbf{Z}^+.$$

We also observe that $g(1) = 0$, $g(2) = 1$, $g(3) = 3$, and $g(4) = 5$, so $g(1) \leq g(2) \leq g(3) \leq g(4)$. Consequently, it appears that g may be a monotone increasing function. The proof that it is monotone increasing is similar to that given for the time-complexity function of binary search. This follows Example 10.49 in Section 10.6, so we leave the details showing that g is monotone increasing to the Section Exercises.

Now with $a = b = 2$ and $c = 1$, the result stated earlier implies that $g \in O(n \log_2 n)$ for *all* $n \in \mathbf{Z}^+$.

Although $n \log_2 n \leq n^2$ for all $n \in \mathbf{Z}^+$, it does *not* follow that because the bubble sort is $O(n^2)$ and the merge sort is $O(n \log_2 n)$, the merge sort is more efficient than the bubble sort for all $n \in \mathbf{Z}^+$. The bubble sort requires less programming effort and generally takes less time than the merge sort for small values of n (depending on factors such as the programming language, the compiler, and the computer). However, as n increases, the ratio of the worst-case running times, as measured by $(cn^2)/(dn \log_2 n) = (c/d)(n/\log_2 n)$, gets arbitrarily large. Consequently, as the input list increases in size, the $O(n^2)$ algorithm (bubble sort) takes significantly more time than the $O(n \log_2 n)$ algorithm (merge sort).

For more on sorting algorithms and their time-complexity functions, the reader should examine [1], [3], [4], [7], and [8] in the chapter references.

EXERCISES 12.3

1. a) Give an example of two lists L_1, L_2, each of which is in ascending order and contains five elements, and where nine comparisons are needed to merge L_1, L_2 by the algorithm given in Lemma 12.1.

b) Let $m, n \in \mathbf{Z}^+$ with $m < n$. Give an example of two lists L_1, L_2, each of which is in ascending order, where L_1 has m elements, L_2 has n elements, and $m + n - 1$ comparisons are needed to merge L_1, L_2 by the algorithm given in Lemma 12.1.

2. Apply the merge sort to each of the following lists. Draw the splitting and merging trees for each application of the procedure.

a) $-1, 0, 2, -2, 3, 6, -3, 5, 1, 4$

b) $-1, 7, 4, 11, 5, -8, 15, -3, -2, 6, 10, 3$

3. Related to the merge sort is a somewhat more efficient procedure called the *quick sort*. Here we start with a list $L: a_1, a_2, \ldots, a_n$, and use a_1 as a pivot to develop two sublists L_1 and L_2 as follows. For $i > 1$, if $a_i < a_1$, place a_i at the end of the first list being developed (this is L_1 at the end of the process); otherwise, place a_i at the end of the second list L_2.

After all a_i, $i > 1$, have been processed, place a_1 at the end of the first list. Now apply quick sort recursively to each of the lists L_1 and L_2 to obtain sublists L_{11}, L_{12}, L_{21}, and L_{22}. Continue the process until each of the resulting sublists contains one element. The sublists are then ordered, and their concatenation gives the ordering sought for the original list L.

Apply quick sort to each list in Exercise 2.

4. Prove that the function g used in the second method to analyze the (worst-case) time-complexity of the merge sort is monotone increasing.

12.4 Weighted Trees and Prefix Codes

Among the topics to which discrete mathematics is applied, coding theory is one wherein different finite structures play a major role. These structures enable us to represent and transmit information that is coded in terms of the symbols in a given alphabet. For instance, the way we most often code, or represent, characters internally in a computer is by means of strings of fixed length, using the symbols 0 and 1.

The codes developed in this section, however, will use strings of different lengths. Why a person should want to develop such a coding scheme and how the scheme can be constructed will be our major concerns in this section.

Suppose we wish to develop a way to represent the letters of the alphabet using strings of 0's and 1's. Since there are 26 letters, we should be able to encode these symbols in terms of sequences of five bits, given that $2^4 < 26 < 2^5$. However, in the English (or any other) language, not all letters occur with the same frequency. Consequently, it would be more efficient to use binary sequences of different lengths, with the most frequently occurring letters (such as e, i, t) represented by the shortest possible sequences. For example, consider $S = \{a, e, n, r, t\}$, a subset of the alphabet. Represent the elements of S by the binary sequences

$$a: 01 \qquad e: 0 \qquad n: 101 \qquad r: 10 \qquad t: 1.$$

If the message "*ata*" is to be transmitted, the binary sequence 01101 is sent. Unfortunately, this sequence is also transmitted for the messages "*etn*", "*atet*", and "*an*".

Consider a second encoding scheme, one given by

$$a: 111 \qquad e: 0 \qquad n: 1100 \qquad r: 1101 \qquad t: 10.$$

Here the message "*ata*" is represented by the sequence 11110111 and there are no other possibilities to confuse the situation. What's more, the labeled complete binary tree shown in Fig. 12.34 can be used to decode the sequence 11110111. Starting at the root, traverse the edge labeled 1 to the right child (of the root). Continuing along the next two edges labeled with 1, we arrive at the leaf labeled a. Hence the unique path from the root to the vertex at a is unambiguously determined by the first three 1's in the sequence 11110111. After we return to the root, the next two symbols in the sequence — namely, 10 — determine the unique path along the edge from the root to its right child, followed by the edge from that child to its left child. This terminates at the vertex labeled t. Again returning to the root, the final three bits of the sequence determine the letter a for a second time. Hence the tree "decodes" 11110111 as *ata*.

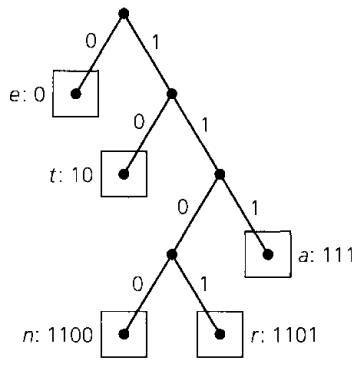


Figure 12.34

Why did the second encoding scheme work out so readily when the first led to ambiguities? In the first scheme, r is represented as 10 and n as 101. If we encounter the symbols 10, how can we determine whether the symbols represent r or the first two symbols of 101, which represent n? The problem is that the sequence for r is a prefix of the sequence for

n. This ambiguity does not occur in the second encoding scheme, suggesting the following definition.

Definition 12.7 A set P of binary sequences (representing a set of symbols) is called a *prefix code* if no sequence in P is the prefix of any other sequence in P.

Consequently, the binary sequences 111, 0, 1100, 1101, 10 constitute a prefix code for the letters a, e, n, r, t, respectively. But how did the complete binary tree of Fig. 12.34 come about? To deal with this problem, we need the following concept.

Definition 12.8 If T is a complete binary tree of height h, then T is called a *full* binary tree if all the leaves in T are at level h.

EXAMPLE 12.17 For the prefix code $P = \{111, 0, 1100, 1101, 10\}$, the longest binary sequence has length 4. Draw the labeled full binary tree of height 4, as shown in Fig. 12.35. The elements of P are assigned to the vertices of this tree as follows. For example, the sequence 10 traces the path from the root r to its right child c_R. Then it continues to the left child of c_R, where the box (marked with the asterisk) indicates completion of the sequence. Returning to the root, the other four sequences are traced out in similar fashion, resulting in the other four boxed vertices. For each boxed vertex remove the subtree (except for the root) that it determines. The resulting pruned tree is the complete binary tree of Fig. 12.34, where no "box" is an ancestor of another "box."

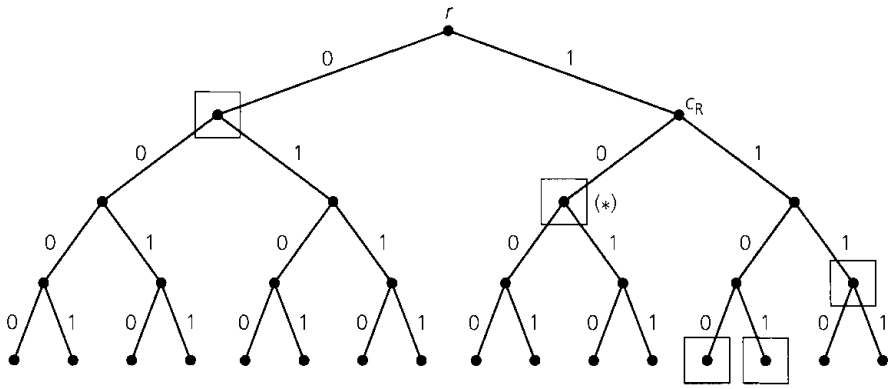


Figure 12.35

We turn now to a method for determining a labeled tree that models a prefix code, where the frequency of occurrence of each symbol in the average text is taken into account — in other words, a prefix code wherein the shorter sequences are used for the more frequently occurring symbols. If there are many symbols, such as all 26 letters of the alphabet, a trial-and-error method for constructing such a tree is not efficient. An elegant construction developed by David A. Huffman (1925–1999) provides a technique for constructing such trees.

The general problem of constructing an efficient tree can be described as follows.

Let $w_1, w_2, \ldots, w_n$ be a set of positive numbers called *weights*, where $w_1 \leq w_2 \leq \cdots \leq w_n$. If $T = (V, E)$ is a complete binary tree with n leaves, assign these weights (in

any one-to-one manner) to the n leaves. The result is called a *complete binary tree for the weights* $w_1, w_2, \ldots, w_n$. The *weight of the tree*, denoted $W(T)$, is defined as $\sum_{i=1}^{n} w_i \ell(w_i)$ where, for each $1 \le i \le n$, $\ell(w_i)$ is the level number of the leaf assigned the weight w_i. The objective is to assign the weights so that $W(T)$ is as small as possible. A complete binary tree T' for these weights is said to be an *optimal* tree if $W(T') \le W(T)$ for any other complete binary tree T for the weights.

Figure 12.36 shows two complete binary trees for the weights 3, 5, 6, and 9. For tree T_1, $W(T_1) = \sum_{i=1}^{4} w_i \ell(w_i) = (3 + 9 + 5 + 6) \cdot 2 = 46$ because each leaf has level number 2. In the case of T_2, $W(T_2) = 3 \cdot 3 + 5 \cdot 3 + 6 \cdot 2 + 9 \cdot 1 = 45$, which we shall find is optimal.

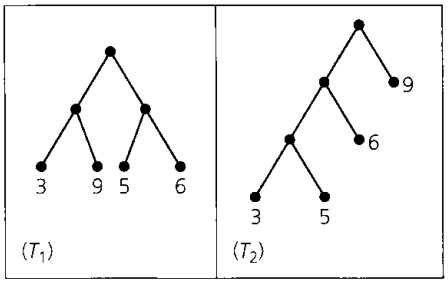


Figure 12.36

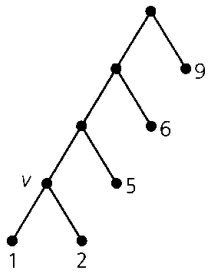


Figure 12.37

The major idea behind Huffman's construction is that in order to obtain an optimal tree T for the n weights $w_1, w_2, w_3, \ldots, w_n$, one considers an optimal tree T' for the $n - 1$ weights $w_1 + w_2, w_3, \ldots, w_n$. (It cannot be assumed that $w_1 + w_2 \le w_3$.) In particular, the tree T' is transformed into T by replacing the leaf v having weight $w_1 + w_2$ by a tree rooted at v of height 1 with left child of weight w_1 and right child of weight w_2. To illustrate, if the tree T_2 in Fig. 12.36 is optimal for the four weights $1 + 2$, 5, 6, 9, then the tree in Fig. 12.37 will be optimal for the five weights 1, 2, 5, 6, 9.

We need the following lemma to establish these claims.

LEMMA 12.2

If T is an optimal tree for the n weights $w_1 \le w_2 \le \cdots \le w_n$, then there exists an optimal tree T' in which the leaves of weights w_1 and w_2 are siblings at the maximal level (in T').

Proof: Let v be an internal vertex of T where the level number of v is maximal for all internal vertices. Let w_x and w_y be the weights assigned to the children x, y of vertex v, with $w_x \le w_y$. By the choice of vertex v, $\ell(w_x) = \ell(w_y) \ge \ell(w_1), \ell(w_2)$. Consider the case of $w_1 < w_x$. (If $w_1 = w_x$, then w_1 and w_x can be interchanged and we would consider the case of $w_2 < w_y$. Applying the following proof to this case, we would find that w_y and w_2 can be interchanged.)

If $\ell(w_x) > \ell(w_1)$, let $\ell(w_x) = \ell(w_1) + j$, for some $j \in \mathbf{Z}^+$. Then $w_1\ell(w_1) + w_x\ell(w_x) = w_1\ell(w_1) + w_x[\ell(w_1) + j] = w_1\ell(w_1) + w_x j + w_x\ell(w_1) > w_1\ell(w_1) + w_1 j + w_x\ell(w_1) = w_1\ell(w_x) + w_x\ell(w_1)$. So $W(T) = w_1\ell(w_1) + w_x\ell(w_x) + \sum_{i \ne 1,x} w_i\ell(w_i) > w_1\ell(w_x) + w_x\ell(w_1) + \sum_{i \ne 1,x} w_i\ell(w_i)$. Consequently, by interchanging the locations of the weights w_1 and w_x, we obtain a tree of smaller weight. But this contradicts the choice of T as an optimal tree. Therefore $\ell(w_x) = \ell(w_1) = \ell(w_y)$. In a similar manner, it can be shown that $\ell(w_y) = \ell(w_2)$, so $\ell(w_x) = \ell(w_y) = \ell(w_1) = \ell(w_2)$. Interchanging the locations of the pair w_1, w_x, and the pair w_2, w_y, we obtain an optimal tree T', where w_1, w_2 are siblings.

From this lemma we see that smaller weights will appear at the higher levels (and thus have higher level numbers) in an optimal tree.

THEOREM 12.8

Let T be an optimal tree for the weights $w_1 + w_2, w_3, \ldots, w_n$, where $w_1 \leq w_2 \leq w_3 \leq \cdots \leq w_n$. At the leaf with weight $w_1 + w_2$ place a (complete) binary tree of height 1 and assign the weights w_1, w_2 to the children (leaves) of this former leaf. The new binary tree T_1 so constructed is then optimal for the weights $w_1, w_2, w_3, \ldots, w_n$.

Proof: Let T_2 be an optimal tree for the weights $w_1, w_2, \ldots, w_n$, where the leaves for weights w_1, w_2 are siblings. Remove the leaves of weights w_1, w_2 and assign the weight $w_1 + w_2$ to their parent (now a leaf). This complete binary tree is denoted T_3 and $W(T_2) = W(T_3) + w_1 + w_2$. Also, $W(T_1) = W(T) + w_1 + w_2$. Since T is optimal, $W(T) \leq W(T_3)$. If $W(T) < W(T_3)$, then $W(T_1) < W(T_2)$, contradicting the choice of T_2 as optimal. Hence $W(T) = W(T_3)$ and, consequently, $W(T_1) = W(T_2)$. So T_1 is optimal for the weights $w_1, w_2, \ldots, w_n$.

Remark. The preceding proof started with an optimal tree T_2 whose existence rests on the fact that there is only a finite number of ways in which we can assign n weights to a complete binary tree with n leaves. Consequently, with a finite number of assignments there is at least one where $W(T)$ is minimal. But finite numbers can be large. This proof establishes the existence of an optimal tree for a set of weights and develops a way for constructing such a tree. To construct such a (Huffman) tree we consider the following algorithm.

Given the n (≥ 2) weights $w_1, w_2, \ldots, w_n$, proceed as follows:

Step 1: Assign the given weights, one each to a set S of n isolated vertices. [Each vertex is the root of a complete binary tree (of height 0) with a weight assigned to it.]

Step 2: While $|S| > 1$ perform the following:

a) Find two trees T, T' in S with the smallest two root weights w, w', respectively.

b) Create the new (complete binary) tree T^* with root weight $w^* = w + w'$ and having T, T' as its left and right subtrees, respectively.

c) Place T^* in S and delete T and T'. [Where $|S| = 1$, the one complete binary tree in S is a Huffman tree.]

We now use this algorithm in the following example.

EXAMPLE 12.18

Construct an optimal prefix code for the symbols a, o, q, u, y, z that occur (in a given sample) with frequencies 20, 28, 4, 17, 12, 7, respectively.

Figure 12.38 shows the construction that follows Huffman's procedure. In part (b) weights 4 and 7 are combined so that we then consider the construction for the weights 11, 12, 17, 20, 28. At each step [in parts (c)–(f) of Fig. 12.38] we create a tree with subtrees rooted at the two smallest weights. These two smallest weights belong to vertices each of which is originally either isolated (a tree with just a root) or the root of a tree obtained earlier in the construction. From the last result, a prefix code is determined as

a: 01 o: 11 q: 1000 u: 00 y: 101 z: 1001.

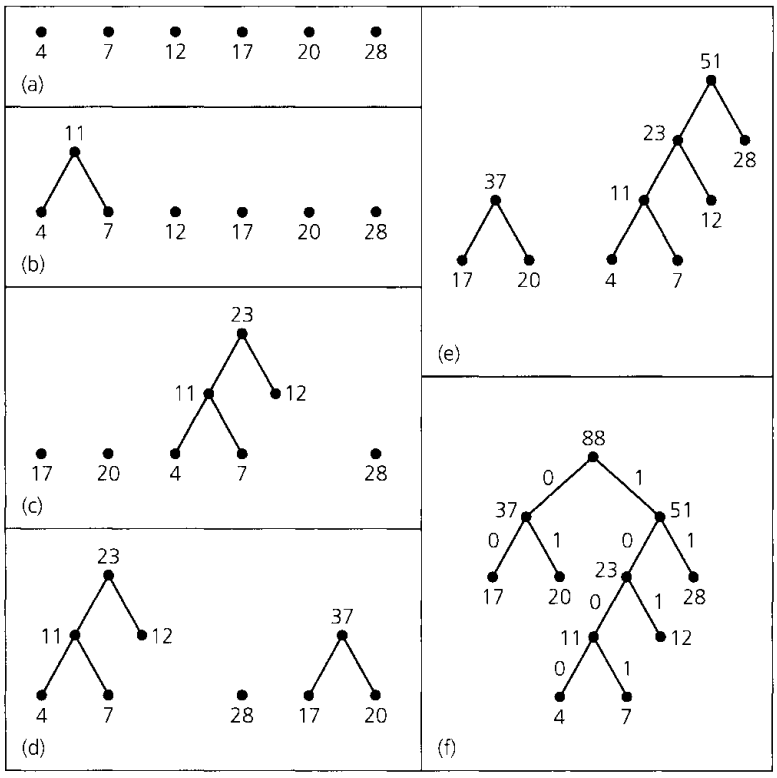


Figure 12.38

Different prefix codes may result from the way the trees T, T' are selected and assigned as the left or right subtree in steps 2(a) and 2(b) in our algorithm and from the assignment of 0 or 1 to the branches (edges) of our final (Huffman) tree.

EXERCISES 12.4

1. For the prefix code given in Fig. 12.34, decode the sequences (a) 1001111101; (b) 10111100110001101; (c) 1101111110010.

2. A code for $\{a, b, c, d, e\}$ is given by a: 00 b: 01 c: 101 d: $x10$ e: $yz1$, where $x, y, z \in \{0, 1\}$. Determine x, y, and z so that the given code is a prefix code.

3. Construct an optimal prefix code for the symbols $a, b, c, \ldots, i, j$ that occur (in a given sample) with respective frequencies 78, 16, 30, 35, 125, 31, 20, 50, 80, 3.

4. How many leaves does a full binary tree have if its height is (a) 3? (b) 7? (c) 12? (d) h?

5. Let $T = (V, E)$ be a complete m-ary tree of height h. This tree is called a *full* m-ary tree if all of its leaves are at level h. If T is a full m-ary tree with height 7 and 279,936 leaves, how many internal vertices are there in T?

6. Let T be a full m-ary tree with height h and v vertices. Determine h in terms of m and v.

7. Using the weights 2, 3, 5, 10, 10, show that the height of a Huffman tree for a given set of weights is not unique. How would you modify the algorithm so as to always produce a Huffman tree of minimal height for the given weights?

8. Let L_i, for $1 \leq i \leq 4$, be four lists of numbers, each sorted in ascending order. The numbers of entries in these lists are 75, 40, 110, and 50, respectively.

a) How many comparisons are needed to merge these four lists by merging L_1 and L_2, merging L_3 and L_4, and then merging the two resulting lists?

b) How many comparisons are needed if we first merge L_1 and L_2, then merge the result with L_3, and finally merge this result with L_4?

c) In order to minimize the total number of comparisons in this merging of the four lists, what order should the merging follow?

d) Extend the result in part (c) to n sorted lists $L_1, L_2, \ldots, L_n$.

12.5 Biconnected Components and Articulation Points

Let $G = (V, E)$ be the loop-free connected undirected graph shown in Fig. 12.39(a), where each vertex represents a communication center. Here an edge $\{x, y\}$ indicates the existence of a communication link between the centers at x and y.

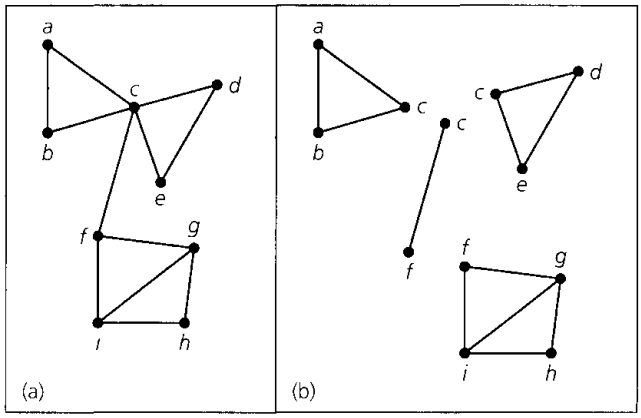


Figure 12.39

By splitting the vertices at c and f, in the suggested fashion, we obtain the collection of subgraphs in part (b) of the figure. These vertices are examples of the following.

Definition 12.9

A vertex v in a loop-free undirected graph $G = (V, E)$ is called an *articulation point* if $\kappa(G - v) > \kappa(G)$; that is, the subgraph $G - v$ has more components than the given graph G.

A loop-free connected undirected graph with no articulation points is called *biconnected.*

A *biconnected component* of a graph is a maximal biconnected subgraph — a biconnected subgraph that is not properly contained in a larger biconnected subgraph.

The graph shown in Fig. 12.39 has the two articulation points, c and f, and its four biconnected components are shown in part (b) of the figure.

In terms of communication centers and links, the articulation points of the graph indicate where the system is most vulnerable. Without articulation points, such a system is more likely to survive disruptions at a communication center, regardless of whether these disruptions are caused by the breakdown of a technical device or by external forces.

The problem of finding the articulation points in a connected graph provides an application for the depth-first spanning tree. The objective here is the development of an algorithm that determines the articulation points of a loop-free connected undirected graph. If no such points exist, then the graph is biconnected. Should such vertices exist, the resulting biconnected components can be used to provide information about such properties as the planarity and chromatic number of the given graph.

The following preliminaries are needed for developing this algorithm.

Returning to Fig. 12.39(a), we see that there are four paths from a to e—namely, (1) $a \rightarrow c \rightarrow e$; (2) $a \rightarrow c \rightarrow d \rightarrow e$; (3) $a \rightarrow b \rightarrow c \rightarrow e$; and (4) $a \rightarrow b \rightarrow c \rightarrow d \rightarrow e$. Now what do these four paths have in common? They all pass through the vertex c, one of the articulation points of G. This observation now motivates our first preliminary result.

LEMMA 12.3

Let $G = (V, E)$ be a loop-free connected undirected graph with $z \in V$. The vertex z is an articulation point of G if and only if there exist distinct $x, y \in V$ with $x \neq z$, $y \neq z$, and such that every path in G connecting x and y contains the vertex z.

Proof: This result follows from Definition 12.9. A proof is requested of the reader in the Section Exercises.

Our next lemma provides an important and useful property of the depth-first spanning tree.

LEMMA 12.4

Let $G = (V, E)$ be a loop-free connected undirected graph with $T = (V, E')$ a depth-first spanning tree of G. If $\{a, b\} \in E$ but $\{a, b\} \notin E'$, then a is either an ancestor or a descendant of b in the tree T.

Proof: From the depth-first spanning tree T, we obtain a preorder listing for the vertices in V. For all $v \in V$, let dfi(v) denote the depth-first index of vertex v—that is, the position of v in the preorder listing. Assume that dfi(a) < dfi(b). Consequently, a is encountered before b in the preorder traversal of T, so a cannot be a descendant of b. If, in addition, vertex a is not an ancestor of b, then b is not in the subtree T_a of T rooted at a. But when we backtrack (through T_a) to a, we find that because $\{a, b\} \in E$, it should have been possible for the depth-first search to go from a to b and to use the edge $\{a, b\}$ in T. This contradiction shows that b is in T_a, so a is an ancestor of b.

If $G = (V, E)$ is a loop-free connected undirected graph, let $T = (V, E')$ be a depth-first spanning tree for G, as shown in Fig. 12.40. By Lemma 12.4, the dotted edge $\{a, b\}$, which is not part of T, indicates an edge that could exist in G. Such an edge is called a *back edge* (relative to T), and here a is an ancestor of b. [Here dfi(a) = 3, whereas dfi(b) = 6.] The dotted edge $\{b, d\}$ in the figure cannot exist in G, also because of Lemma 12.4. Thus all edges of G are either edges in T or back edges (relative to T).

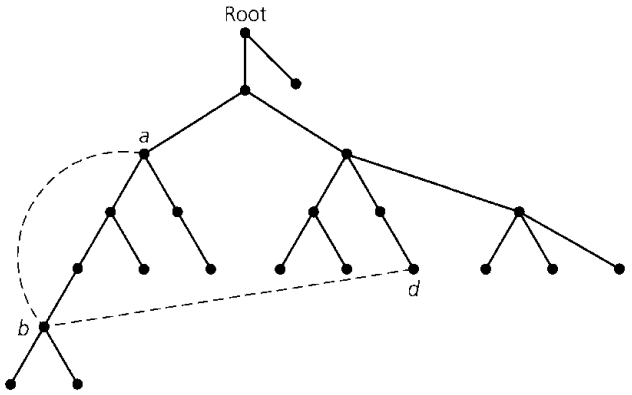


Figure 12.40

Our next example provides further insight into the relationship between the articulation points of a graph G and a depth-first spanning tree of G.

EXAMPLE 12.19

In part (1) of Fig. 12.41 we have a loop-free connected undirected graph $G = (V, E)$. Applying Lemma 12.3 to vertex a, for example, we find that the only path in G from b to i passes through a. In the case of vertex d, we apply the same lemma and consider the vertices a and h. Now we find that although there are four paths from a to h, all four pass through vertex d. Consequently, vertices a and d are two of the articulation points in G. The vertex h is the only other articulation point. Can you find two vertices in G for which all connecting paths (for these vertices) in G pass through h?

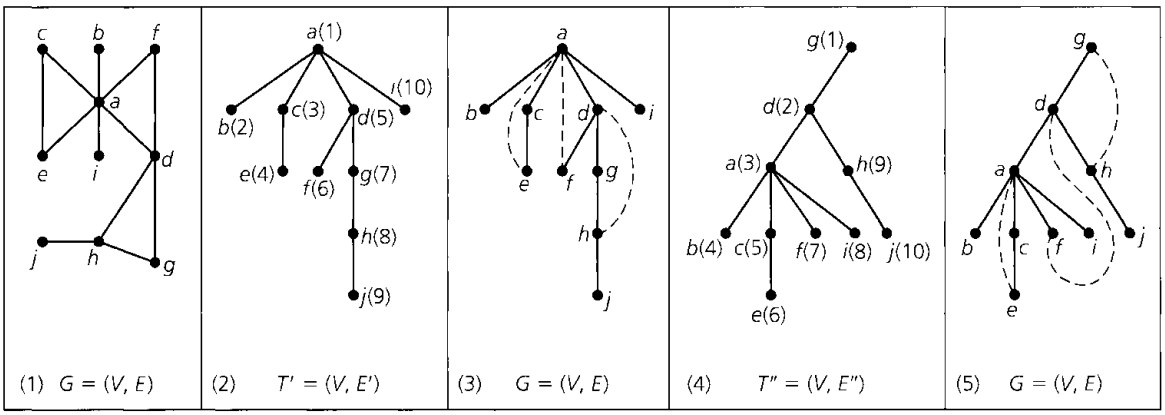


Figure 12.41

Applying the depth-first search algorithm, with the vertices of G ordered alphabetically, in part (2) of Fig. 12.41, we find the depth-first spanning tree $T' = (V, E')$ for G, where a has been chosen as the root. The parenthesized integer next to each vertex indicates the order in which that vertex is visited during the prescribed depth-first search. Part (3) of the figure incorporates the three back edges (relative to T, in G) that are missing from part (2).

For the tree T', the root a, which is an articulation point in G, has more than one child. The articulation point d has a child — namely, g — with no back edge from g or any of its descendants (h and j) to an ancestor of d [as we see in part (3) of Fig. 12.41]. The same is true for the articulation point h. Its child j has (no children and) no back edge to an ancestor of h.

In part (4) of the figure, $T'' = (V, E'')$ is the depth-first spanning tree for the vertices ordered alphabetically once again, but this time vertex g has been chosen as the root. As in part (2) of the figure, the parenthesized integer next to each vertex indicates the order in which that vertex is visited during this depth-first search. The three back edges (relative to T'', in G) that are missing from T'' are shown in part (5) of the figure.

The root g of T'' has only one child and g is not an articulation point in G. Further, for each of the articulation points there is at least one child with no back edge from that child or one of its descendants to an ancestor of the articulation point. To be more specific, from part (5) of Fig. 12.41 we find that for the articulation point a we may use any of the children b, c or i, but not f; for d that child is a; and for h the child is j.

The observations made in Example 12.19 now lead us to the following.

LEMMA 12.5

Let $G = (V, E)$ be a loop-free connected undirected graph with $T = (V, E')$ a depth-first spanning tree of G. If r is the root of T, then r is an articulation point of G if and only if r has at least two children in T.

Proof: If r has only one child — say, c — then all the other vertices of G are descendants of c (and r) in T. So if x, y are two distinct vertices of T, neither of which is r, then in the subtree T_c, rooted at c, there is a path from x to y. Since r is not a vertex in T_c, r is not on this path. Consequently, r is not an articulation point in G — by virtue of Lemma 12.3. Conversely, let r be the root of the depth-first spanning tree T and let c_1, c_2 be children of r. Let x be a vertex in T_{c_1}, the subtree of T rooted at c_1. Similarly, let y be a vertex in T_{c_2}, the subtree of T rooted at c_2. Could there be a path from x to y in G that avoids r? If so, there is an edge $\{v_1, v_2\}$ in G with v_1 in T_{c_1} and v_2 in T_{c_2}. But this contradicts Lemma 12.4.

Our final preliminary result settles the issue of when a vertex, that is not the root of a depth-first spanning tree, is an articulation point of a graph.

LEMMA 12.6

Let $G = (V, E)$ be a loop-free connected undirected graph with $T = (V, E')$ a depth-first spanning tree for G. Let r be the root of T and let $v \in V$, $v \neq r$. Then v is an articulation point of G if and only if there exists a child c of v with no back edge (relative to T, in G) from a vertex in T_c, the subtree rooted at c, to an ancestor of v.

Proof: Suppose that vertex v has a child c such that there is no back edge (relative to T, in G) from a vertex in T_c to an ancestor of v. Then every path (in G) from r to c passes through v. From Lemma 12.3 it then follows that v is an articulation point of G.

To establish the converse, let the nonroot vertex v of T satisfy the following: For each child c of v there is a back edge (relative to T, in G) from a vertex in T_c, the subtree rooted at c, to an ancestor of v. Now let $x, y \in V$ with $x \neq v$, $y \neq v$. We consider the following three possibilities:

1) If neither x nor y is a descendant of v, as in part (1) of Fig. 12.42, delete from T the subtree T_v rooted at v. The resulting subtree (of T) contains x, y and a path from x to y that does not pass through v, so v is not an articulation point of G.

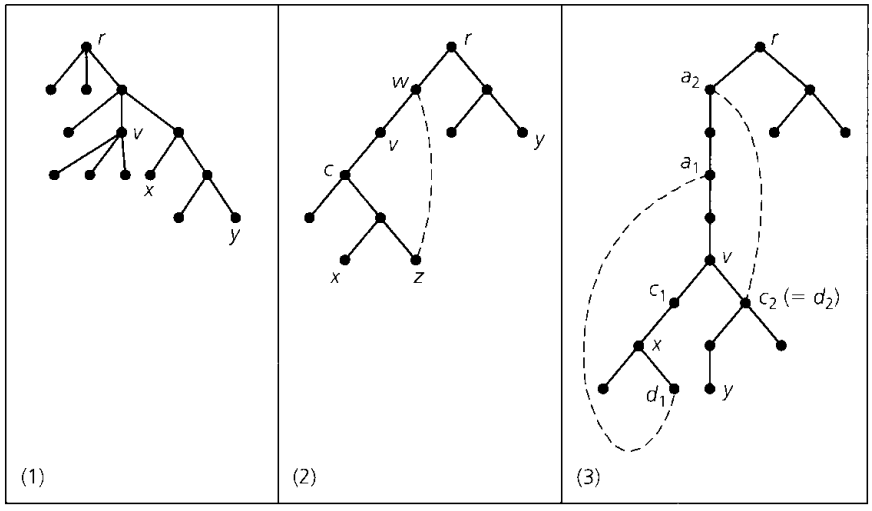


Figure 12.42

2) If one of x, y — say, x — is a descendant of v but y is not, then x is a child of v or a descendant of a child c of v [as in part (2) of Fig. 12.42]. From the hypothesis there is a back edge (relative to T, in G) from some $z \in T_c$ to an ancestor w of v. Since $x, z \in T_c$, there is a path p_1 from x to z (that does not pass through v). Then, as neither w nor y is a descendant of v, from part (1) there is a path p_2 from w to y that does not pass through v. The edges in p_1, p_2 together with the edge $\{z, w\}$ provide a path from x to y that does not pass through v — and once again, v is not an articulation point.

3) Finally, suppose that both x, y are descendants of v, as in part (3) of Fig. 12.42. Here c_1, c_2 are children of v — perhaps, with $c_1 = c_2$ — and x is a vertex in T_{c_1}, the subtree rooted at c_1, while y is a vertex in T_{c_2}, the subtree rooted at c_2. From the hypothesis, there exist back edges $\{d_1, a_1\}$ and $\{d_2, a_2\}$ (relative to T, in G), where d_1, d_2 are descendants of v and a_1, a_2 are ancestors of v. Further, there is a path p_1 from x to d_1 in T_{c_1} and a path p_2 from y to d_2 in T_{c_2}. As neither a_1 nor a_2 is a descendant of v, from part (1) we have a path p (in T) from a_1 to a_2, where p avoids v. Now we can do the following: (i) Go from x to d_1 using path p_1; (ii) Go from d_1 to a_1 on the edge $\{d_1, a_1\}$; (iii) Continue to a_2 using path p; (iv) Go from a_2 to d_2 on the edge $\{a_2, d_2\}$; and (v) Finish at y using the path p_2 from d_2 to y. This provides a path from x to y that avoids v so v is not an articulation point of G and this completes the proof.

Using the results from the preceding four lemmas, we once again start with a loop-free connected undirected graph $G = (V, E)$ with depth-first spanning tree T. For $v \in V$, where v is not the root of T, we let $T_{v,c}$ be the subtree consisting of edge $\{v, c\}$ (c a child of v) together with the tree T_c rooted at c. If there is no back edge from a descendant of v in $T_{v,c}$ to an ancestor of v (and v has at least one ancestor — the root of T), then the splitting of vertex v results in the separation of $T_{v,c}$ from G, and v is an articulation point. If no other articulation points of G occur in $T_{v,c}$, then the addition to $T_{v,c}$ of all other edges in G determined by the vertices in $T_{v,c}$ (the subgraph of G induced by the vertices in $T_{v,c}$) results in a biconnected component of G. A root has no ancestors, and it is an articulation point if and only if it has more than one child.

The depth-first spanning tree preorders the vertices of G. For $x \in V$ let $\text{dfi}(x)$ denote the depth-first index of x in that preorder. If y is a descendant of x, then $\text{dfi}(x) < \text{dfi}(y)$. For y an ancestor of x, $\text{dfi}(x) > \text{dfi}(y)$. Define $\text{low}(x) = \min\{\text{dfi}(y) | y \text{ is adjacent in } G \text{ to either } x \text{ or a descendant of } x\}$. If z is the parent of x (in T), then there are two possibilities to consider:

1) $\text{low}(x) = \text{dfi}(z)$: In this case T_x, the subtree rooted at x, contains no vertex that is adjacent to an ancestor of z by means of a back edge of T. Hence z is an articulation point of G. If T_x contains no articulation points, then T_x together with edge $\{z, x\}$ spans a biconnected component of G (that is, the subgraph of G induced by vertex z and the vertices in T_x is a biconnected component of G). Now remove T_x and the edge $\{z, x\}$ from T, and apply this idea to the remaining subtree of T.

2) $\text{low}(x) < \text{dfi}(z)$: Here there is a descendant of z in T_x that is joined [by a back edge (relative to T, in G)] to an ancestor of z.

To deal in an efficient manner with these ideas, we develop the following algorithm. Let $G = (V, E)$ be a loop-free connected undirected graph.

Step 1: Find the depth-first spanning tree T for G according to a prescribed order. Let $x_1, x_2, \ldots, x_n$ be the vertices of G preordered by T. Then $\text{dfi}(x_j) = j$ for all $1 \leq j \leq n$.

Step 2: Start with x_n and continue back to $x_{n-1}, x_{n-2}, \ldots, x_3, x_2, x_1$, determining $\text{low}(x_j)$, for $j = n, n-1, n-2, \ldots, 3, 2, 1$, recursively, as follows:

a) $\text{low}'(x_j) = \min\{\text{dfi}(z) | z \text{ is adjacent in } G \text{ to } x_j\}$.

b) If $c_1, c_2, \ldots, c_m$ are the children of x_j, then $\text{low}(x_j) = \min\{\text{low}'(x_j), \text{low}(c_1), \text{low}(c_2), \ldots, \text{low}(c_m)\}$. [No problem arises here, for the vertices are examined in the reverse order to the given preorder. Consequently, if c is a child of p, then $\text{low}(c)$ is determined before $\text{low}(p)$.]

Step 3: Let w_j be the parent of x_j in T. If $\text{low}(x_j) = \text{dfi}(w_j)$, then w_j is an articulation point of G, unless w_j is the root of T and w_j has no child in T other than x_j. Moreover, in either situation the subtree rooted at x_j together with the edge $\{w_j, x_j\}$ is part of a biconnected component of G.

EXAMPLE 12.20

We apply this algorithm to the graph $G = (V, E)$ shown in part (i) of Fig. 12.43.

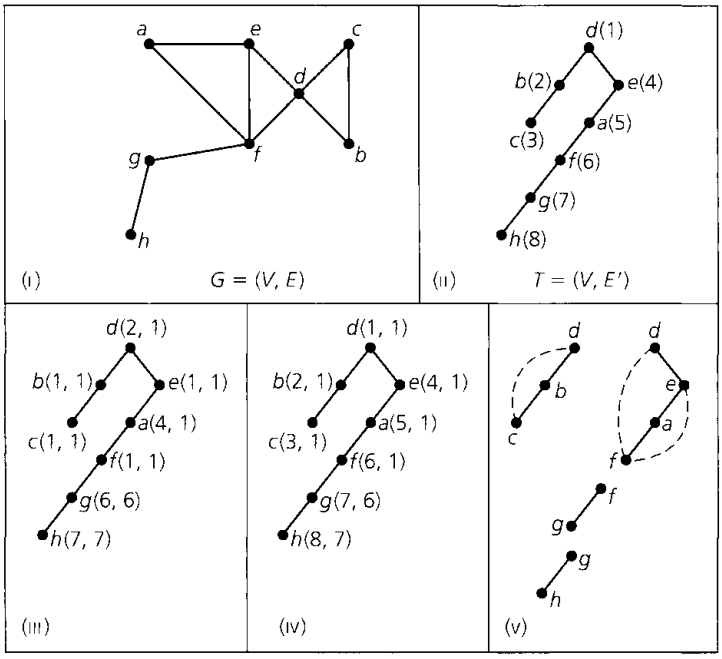


Figure 12.43

In part (ii) of the figure we have the depth-first spanning tree $T = (V, E')$ for G with d as the root. (Here the order followed for the vertices of G is alphabetic.) Next to each vertex v of T [in part (ii)] is the $\text{dfi}(v)$. These labels tell us the order in which the vertices of G are first visited.

For step (2) of the algorithm we go in the reverse order from the depth-first search and start with vertex $h (= x_8)$. Since $\{g, h\} \in E$ and h is not adjacent to any other vertex of G we have $\text{low}'(h) = \text{dfi}(g)$ [$= \text{dfi}(x_7)$] $= 7$. Further, as h has no children, it follows that $\text{low}(h) = \text{low}'(h) = 7$. This accounts for the label (7, 7) [$= (\text{low}'(h), \text{low}(h))$] next

to h in part (iii) of Fig. 12.43. Continuing next with g, and then f, we obtain the labels (6, 6) for g, and (1, 1) for f, since $\text{low}'(g) = \text{low}(g) = 6$ and $\text{low}'(f) = \text{low}(f) = 1$. Since $\{a, e\}, \{a, f\} \in E$ with $\text{dfi}(e) = 4$ and $\text{dfi}(f) = 6$, for vertex a we have $\text{low}'(a) = \min\{4, 6\} = 4$. Then we find that $\text{low}(a) = \min\{4, \text{low}(f)\} = \min\{4, 1\} = 1$. Hence the label (4, 1) for vertex a. Continuing back through e, c, b, and d, we obtain the labels $(\text{low}'(x_i), \text{low}(x_i))$ for $i = 4, 3, 2, 1$. Consequently, by applying step (2) of the algorithm we arrive at the tree in Fig. 12.43 (iii).

In part (iv) of Fig. 12.43 the ordered pair next to each vertex v is $(\text{dfi}(v), \text{low}(v))$. Applying step (3) of the algorithm to the tree in part (iv), at this point we go in reverse order once again. First we deal with vertex h $(= x_8)$. Since g is the parent of h (in T) and $\text{low}(h) = 7 = \text{dfi}(g)$, g is an articulation point of G and the edge $\{h, g\}$ is a biconnected component of G. Deleting the subtree rooted at g from T, we continue with vertex g $(= x_7)$. Here f is the parent of g (in the tree $T - h$) and $\text{low}(g) = 6 = \text{dfi}(f)$, so f is another articulation point — with edge $\{g, f\}$ the corresponding biconnected component.

Continuing now with the tree $(T - h) - g$, as we go from f to a to e, and then from c to b, we find no new articulation points among the four vertices a, e, c, and b. Since vertex d is the root of T and d has two children — namely, the vertices b and e, it then follows from Lemma 12.5 that d is an articulation point of G. The vertices d, e, a, f induce the biconnected component consisting of the tree edges $\{f, a\}, \{a, e\}, \{e, d\}$ and the back edges (relative to T, in G) $\{f, e\}$ and $\{f, d\}$. Finally, the cycle induced (in G) by the vertices b, c and d provides the fourth biconnected component.

Part (v) of Fig. 12.43 shows the three articulation points g, f, and d, and the four biconnected components of G.

EXERCISES 12.5

1. Find the articulation points and biconnected components for the graph shown in Fig. 12.44.

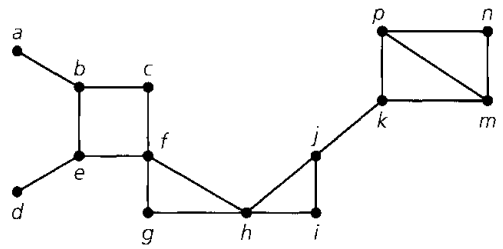


Figure 12.44

2. Prove Lemma 12.3.

3. Let $T = (V, E)$ be a tree with $|V| = n \geq 3$.

a) What are the smallest and the largest numbers of articulation points that T can have? Describe the trees for each of these cases.

b) How many biconnected components does T have in each of the cases in part (a)?

4. a) Let $T = (V, E)$ be a tree. If $v \in V$, prove that v is an articulation point of T if and only if $\deg(v) > 1$.

b) Let $G = (V, E)$ be a loop-free connected undirected graph with $|E| \geq 1$. Prove that G has at least two vertices that are not articulation points.

5. If $B_1, B_2, \ldots, B_k$ are the biconnected components of a loop-free connected undirected graph G, how is $\chi(G)$ related to $\chi(B_i)$, $1 \leq i \leq k$? [Recall that $\chi(G)$ denotes the chromatic number of G, as defined in Section 11.6.]

6. Let $G = (V, E)$ be a loop-free connected undirected graph with biconnected components $B_1, B_2, \ldots, B_8$. For $1 \leq i \leq 8$, the number of distinct spanning trees for B_i is n_i. How many distinct spanning trees exist for G?

7. Let $G = (V, E)$ be a loop-free connected undirected graph with $|V| \geq 3$. If G has no articulation points, prove that G has no pendant vertices.

8. For the loop-free connected undirected graph G in Fig. 12.43(i), order the vertices alphabetically.

a) Determine the depth-first spanning tree T for G with e as the root.

b) Apply the algorithm developed in this section to the tree T in part (a) to find the articulation points and biconnected components of G.

9. Answer the questions posed in the previous exercise but this time order the vertices as h, g, f, e, d, c, b, a and let c be the root of T.

10. Let $G = (V, E)$ be a loop-free connected undirected graph, where $V = \{a, b, c, \ldots, h, i, j\}$. Ordering the vertices alphabetically, the depth-first spanning tree T for G — with a as the root — is given in Fig. 12.45(i). In part (ii) of the figure the ordered pair next to each vertex v provides $(\text{low}'(v), \text{low}(v))$. Determine the articulation points and the spanning trees for the biconnected components of G.

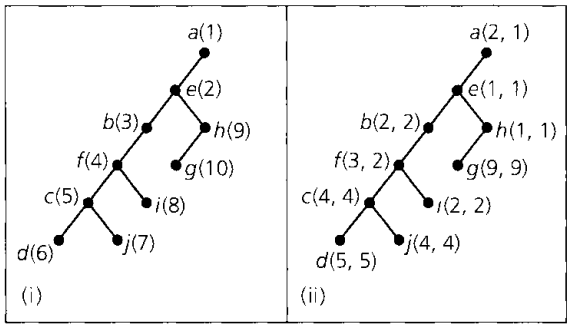


Figure 12.45

11. In step (2) of the algorithm for articulation points, is it really necessary to compute $\text{low}(x_1)$ and $\text{low}(x_2)$?

12. Let $G = (V, E)$ be a loop-free connected undirected graph with $v \in V$.

a) Prove that $\overline{G - v} = \overline{G} - v$.

b) If v is an articulation point of G, prove that v cannot be an articulation point of $\overline{G}$.

13. If $G = (V, E)$ is a loop-free undirected graph, we call G *color-critical* if $\chi(G - v) < \chi(G)$ for all $v \in V$. (We examined such graphs earlier, in Exercise 19 of Section 11.6.) Prove that a color-critical graph has no articulation points.

14. Does the result in Lemma 12.4 remain true if $T = (V, E')$ is a breadth-first spanning tree for $G = (V, E)$?

12.6 Summary and Historical Review

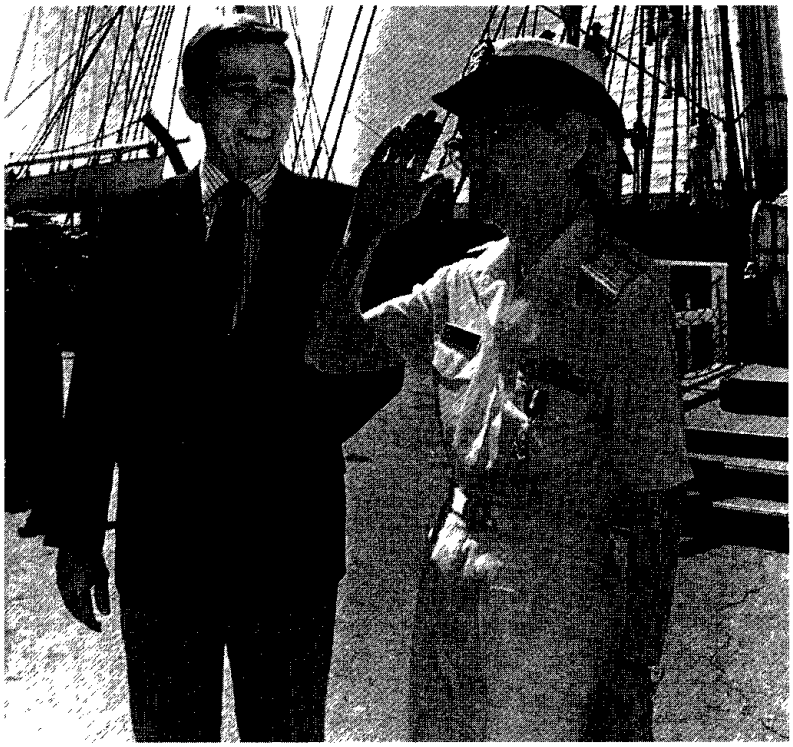
The structure now called a tree first appeared in 1847 in the work of Gustav Kirchhoff (1824–1887) on electrical networks. The concept also appeared at this time in *Geometrie die Lage*, by Karl von Staudt (1798–1867). In 1857 trees were rediscovered by Arthur Cayley (1821–1895), who was unaware of these earlier developments. The first to call the structure a "tree," Cayley used it in applications dealing with chemical isomers. He also investigated the enumeration of certain classes of trees. In his first work on trees, Cayley enumerated unlabeled rooted trees. This was then followed by the enumeration of unlabeled ordered trees. Two of Cayley's contemporaries who also studied trees were Carl Borchardt (1817–1880) and Marie Ennemond Jordan (1838–1922).

Arthur Cayley (1821–1895)

The formula n^{n-2} for the number of labeled trees on n vertices (Exercise 21 at the end of Section 12.1) was discovered in 1860 by Carl Borchardt. Cayley later gave an independent development of the formula, in 1889. Since then, there have been other derivations. These are surveyed in the book by J. W. Moon [10].

The paper by G. Polya [11] is a pioneering work on the enumeration of trees and other combinatorial structures. Polya's theory of enumeration, which we shall see in Chapter 16, was developed in this work. For more on the enumeration of trees, the reader should see Chapter 15 of F. Harary [5]. The article by D. R. Shier [12] provides a labyrinth of several different techniques for calculating the number of spanning trees for $K_{2,n}$.

The high-speed digital computer has proved to be a constant impetus for the discovery of new applications of trees. The first application of these structures was in the manipulation of algebraic formulae. This dates back to 1951 in the work of Grace Murray Hopper. Since then, computer applications of trees have been widely investigated. In the beginning, particular results appeared only in the documentation of specific algorithms. The first general survey of the applications of trees was made in 1961 by Kenneth Iverson as part of a broader survey on data structures. Such ideas as preorder and postorder can be traced to the early 1960s, as evidenced in the work of Zdzislaw Pawlak, Lyle Johnson, and Kenneth Iverson. At this time Kenneth Iverson also introduced the name and the notation, namely $\lceil x \rceil$, for the ceiling of a real number x. Additional material on these orders and the procedures for their implementation on a computer can be found in Chapter 3 of the text by A. V. Aho, J. E. Hopcroft, and J. D. Ullman [1]. In the article by J. E. Atkins, J. S. Dierckman, and K. O'Bryant [2], the notion of preorder is used to develop an optimal route for snow removal.



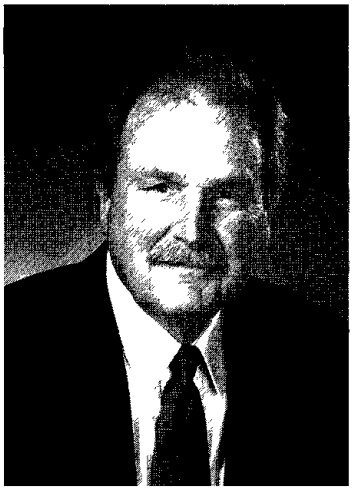
Rear Admiral Grace Murray Hopper (1906–1992) salutes as she and Navy Secretary John Lehman leave the U.S.S *Constitution*.

AP/World Wide Photos

If $G = (V, E)$ is a loop-free undirected graph, then the depth-first search and the breadth-first search (given in Section 12.2) provide ways to determine whether the given graph is connected. The algorithms developed for these searching procedures are also important in developing other algorithms. For example, the depth-first search arises in the algorithm for finding the articulation points and biconnected components of a loop-free connected undirected graph. If $|V| = n$ and $|E| = e$, then it can be shown that both the depth-first search and the breadth-first search have time-complexity $O(\max\{n, e\})$. For most graphs $e > n$, so the algorithms are generally considered to have time-complexity $O(e)$. These ideas are developed in great detail in Chapter 7 of S. Baase and A. Van Gelder [3], where the coverage also includes an analysis of the time-complexity function for the algorithm (of Section 12.5) that determines articulation points (and biconnected components). Chapter 6 of the text by A. V. Aho, J. E. Hopcroft, and J. D. Ullman [1] also deals with the depth-first search, whereas Chapter 7 covers the breadth-first search and the algorithm for articulation points.

More on the properties and computer applications of trees is given in Section 3 of Chapter 2 in the work by D. E. Knuth [7]. Sorting techniques and their use of trees can be further studied in Chapter 11 of A. V. Aho, J. E. Hopcroft, and J. D. Ullman [1] and in Chapter 7 of T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein [4]. An extensive investigation will warrant the coverage found in the text by D. E. Knuth [8].

The technique in Section 12.4 for designing prefix codes is based on methods developed by D. A. Huffman [6].



David A. Huffman

University of Florida, Department of Computer and Information Science and Engineering

Finally, Chapter 7 of C. L. Liu [9] deals with trees, cycles, cut-sets, and the vector spaces associated with these ideas. The reader with a background in linear or abstract algebra should find this material of interest.

REFERENCES

1. Aho, Alfred V., Hopcroft, John E., and Ullman, Jeffrey D. *Data Structures and Algorithms.* Reading, Mass.: Addison-Wesley, 1983.
2. Atkins, Joel E., Dierckman, Jeffrey S., and O'Bryant, Kevin. "A Real Snow Job." *The UMAP Journal*, Fall no. 3 (1990): pp. 231–239.

3. Baase, Sara, and Van Gelder, Allen. *Computer Algorithms: Introduction to Design and Analysis*, 3rd ed. Reading, Mass.: Addison-Wesley, 2000.
4. Cormen, Thomas H., Leiserson, Charles E., Rivest, Ronald L., and Stein, Clifford. *Introduction to Algorithms*, 2nd ed. Boston, Mass.: McGraw-Hill, 2001.
5. Harary, Frank. *Graph Theory*. Reading, Mass.: Addison-Wesley, 1969.
6. Huffman, David A. "A Method for the Construction of Minimum Redundancy Codes." *Proceedings of the IRE* 40 (1952): pp. 1098–1101.
7. Knuth, Donald E. *The Art of Computer Programming*, Vol. 1, 2nd ed. Reading, Mass.: Addison-Wesley, 1973.
8. Knuth, Donald E. *The Art of Computer Programming*, Vol. 3. Reading, Mass.: Addison-Wesley, 1973.
9. Liu, C. L. *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill, 1968.
10. Moon, John Wesley. *Counting Labelled Trees*. Canadian Mathematical Congress, Montreal, Canada, 1970.
11. Polya, George. "Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und Chemische Verbindungen." *Acta Mathematica* 68 (1937): pp. 145–234.
12. Shier, Douglas R. "Spanning Trees: Let Me Count the Ways." *Mathematics Magazine* 73 (2000): pp. 376–381.

SUPPLEMENTARY EXERCISES

1. Let $G = (V, E)$ be a loop-free undirected graph with $|V| = n$. Prove that G is a tree if and only if $P(G, \lambda) = \lambda(\lambda - 1)^{n-1}$.

2. A telephone communication system is set up at a company where 125 executives are employed. The system is initialized by the president, who calls her four vice presidents. Each vice president then calls four other executives, some of whom in turn call four others, and so on. (Each executive who does make a call will actually make four calls.)

a) How many calls are made in reaching all 125 executives?

b) How many executives, aside from the president, are required to make calls?

3. Let T be a complete binary tree with the vertices of T ordered by a preorder traversal. This traversal assigns the label 1 to all internal vertices of T and the label 0 to each leaf. The sequence of 0's and 1's that results from the preorder traversal of T is called the tree's *characteristic sequence*.

a) Find the characteristic sequence for the complete binary tree shown in Fig. 12.17.

b) Determine the complete binary trees for the characteristic sequences

i) 1011001010100 and

ii) 1011110000101011000.

c) What are the last two symbols in the characteristic sequence for all complete binary trees? Why?

4. For $k \in \mathbf{Z}^+$, let $n = 2^k$, and consider the list $L: a_1, a_2, a_3, \ldots, a_n$. To sort L in ascending order, first compare the entries a_i and $a_{i+(n/2)}$, for each $1 \leq i \leq n/2$. For the resulting 2^{k-1} ordered pairs, merge sort the ith and $(i + (n/4))$-th ordered pairs, for each $1 \leq i \leq n/4$. Now do a merge sort on the ith and $(i + (n/8))$-th ordered quadruples, for each $1 \leq i \leq n/8$. Continue the process until the elements of L are in ascending order.

a) Apply this sorting procedure to the list

$$L: 11, 3, 4, 6, -5, 7, 35, -2, 1, 23, 9, 15, 18, 2, -10, 5.$$

b) If $n = 2^k$, how many comparisons at most does this procedure require?

5. Let $G = (V, E)$ be a loop-free undirected graph. If $\deg(v) \geq 2$ for all $v \in V$, prove that G contains a cycle.

6. Let $T = (V, E)$ be a rooted tree with root r. Define the relation $\mathcal{R}$ on V by $x \mathcal{R} y$, for $x, y \in V$, if $x = y$ or if x is on the path from r to y. Prove that $\mathcal{R}$ is a partial order.

7. Let $T = (V, E)$ be a tree with $V = \{v_1, v_2, \ldots, v_n\}$, for $n \geq 2$. Prove that the number of pendant vertices in T is equal to

$$2 + \sum_{\deg(v_i) \geq 3} (\deg(v_i) - 2).$$

8. Let $G = (V, E)$ be a loop-free undirected graph. Define the relation $\mathcal{R}$ on E as follows: If $e_1, e_2 \in E$, then $e_1 \mathcal{R} e_2$ if $e_1 = e_2$ or if e_1 and e_2 are edges of a cycle C in G.

a) Verify that $\mathcal{R}$ is an equivalence relation on E.

b) Describe the partition of E induced by $\mathcal{R}$.

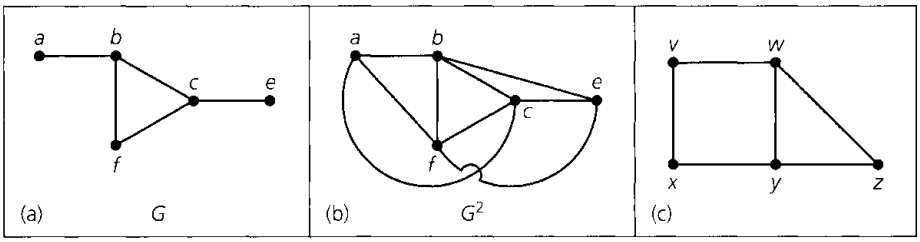


Figure 12.46

9. If $G = (V, E)$ is a loop-free connected undirected graph and $a, b \in V$, then we define the *distance* from a to b (or from b to a), denoted $d(a, b)$, as the length of a shortest path (in G) connecting a and b. (This is the number of edges in a shortest path connecting a and b and is 0 when $a = b$.)

For any loop-free connected undirected graph $G = (V, E)$, the *square* of G, denoted G^2, is the graph with vertex set V (the same as G) and edge set defined as follows: For distinct $a, b \in V$, $\{a, b\}$ is an edge in G^2 if $d(a, b) \leq 2$ (in G). In parts (a) and (b) of Fig. 12.46, we have a graph G and its square.

a) Find the square of the graph in part (c) of the figure.

b) Find G^2 if G is the graph $K_{1,3}$.

c) If G is the graph $K_{1,n}$, for $n \geq 4$, how many edges are added to G in order to construct G^2?

d) For any loop-free connected undirected graph G, prove that G^2 has no articulation points.

10. a) Let $T = (V, E)$ be a complete 6-ary tree of height 8. If T is balanced, but not full, determine the minimum and maximum values for $|V|$.

b) Answer part (a) if $T = (V, E)$ is a complete m-ary tree of height h.

11. The rooted *Fibonacci trees* T_n, $n \geq 1$, are defined recursively as follows:

1) T_1 is the rooted tree consisting of only the root;

2) T_2 is the same as T_1 — it too is a rooted tree that consists of a single vertex; and

3) For $n \geq 3$, T_n is the rooted binary tree with T_{n-1} as its left subtree and T_{n-2} as its right subtree.

The first six rooted Fibonacci trees are shown in Fig. 12.47:

a) For $n \geq 1$, let ℓ_n count the number of leaves in T_n. Find and solve a recurrence relation for ℓ_n.

b) Let i_n count the number of internal vertices for the tree T_n, where $n \geq 1$. Find and solve a recurrence relation for i_n.

c) Determine a formula for v_n, the total number of vertices in T_n, where $n \geq 1$.

12. a) The graph in part (a) of Fig. 12.48 has exactly one spanning tree — namely, the graph itself. The graph in Fig. 12.48(b) has four nonidentical, though isomorphic, spanning trees. In part (c) of the figure we find three of the nonidentical spanning trees for the graph in part (d). Note that T_2 and T_3 are isomorphic, but T_1 is not isomorphic to T_2 (or T_3). How many nonidentical spanning trees exist for the graph in Fig. 12.48(d)?

b) In Fig. 12.48(e) we generalize the graphs in parts (a), (b), and (d) of the figure. For each $n \in \mathbf{Z}^+$, the graph G_n is $K_{2,n}$.

If t_n counts the number of nonidentical spanning trees for G_n, find and solve a recurrence relation for t_n.

13. Let $G = (V, E)$ be the undirected connected "ladder graph" shown in Fig. 12.49. For $n \geq 0$, let a_n count the number of spanning trees of G, whereas b_n counts the number of these spanning trees that contain the edge $\{x_1, y_1\}$.

a) Explain why $a_n = a_{n-1} + b_n$.

b) Find an equation that expresses b_n in terms of a_{n-1} and b_{n-1}.

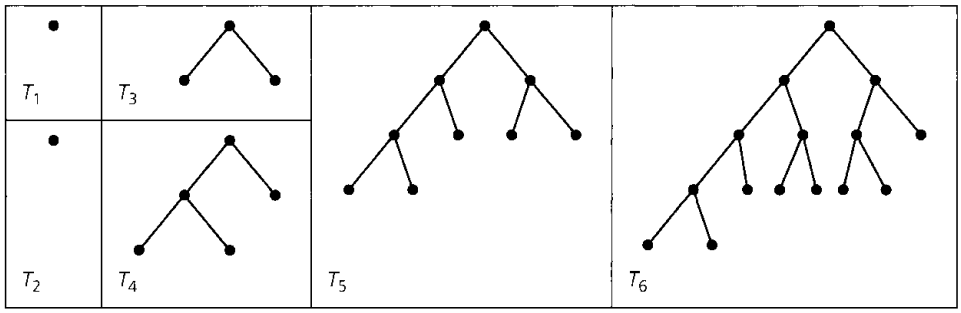


Figure 12.47

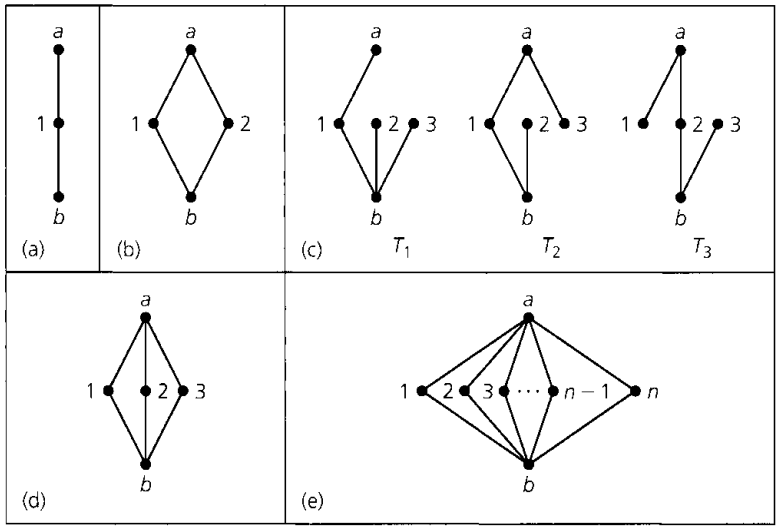


Figure 12.48

c) Use the results in parts (a) and (b) to set up and solve a recurrence relation for a_n.

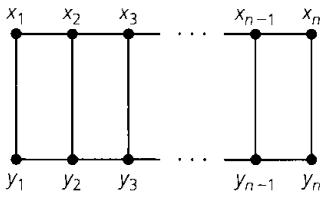


Figure 12.49

14. Let $T = (V, E)$ be a tree where $|V| = v$ and $|E| = e$. The tree T is called *graceful* if it is possible to assign the labels $\{1, 2, 3, \ldots, v\}$ to the vertices of T in such a manner that the induced edge labeling — where each edge $\{i, j\}$ is assigned the label $|i - j|$, for $i, j \in \{1, 2, 3, \ldots, v\}$, $i \neq j$ — results in the e edges being labeled by $1, 2, 3, \ldots, e$.

a) Prove that every path on n vertices, $n \geq 2$, is graceful.

b) For $n \in \mathbf{Z}^+$, $n \geq 2$, show that $K_{1,n}$ is graceful.

c) If $T = (V, E)$ is a tree with $4 \leq |V| \leq 6$, show that T is graceful. (It has been conjectured that every tree is graceful.)

15. For an undirected graph $G = (V, E)$ a subset of I of V is called *independent* when no two vertices in I are adjacent. If, in addition, $I \cup \{x\}$ is not independent for each $x \in V - I$, then we say that I is a *maximal independent* set (of vertices).

The two graphs in Fig. 12.50 are examples of special kinds of trees called caterpillars. In general, a tree $T = (V, E)$ is a *caterpillar* when there is a (maximal) path p such that, for all $v \in V$, either v is on the path p or v is adjacent to a vertex on the path p. This path p is called the *spine* of the caterpillar.

a) How many maximal independent sets of vertices are there for each of the caterpillars in parts (i) and (ii) of Fig. 12.50?

b) For $n \in \mathbf{Z}^+$, with $n \geq 3$, let a_n count the number of maximal independent sets in a caterpillar T whose spine contains n vertices. Find and solve a recurrence relation for a_n. [The reader may wish to reexamine part (a) of Supplementary Exercise 21 in Chapter 11.]

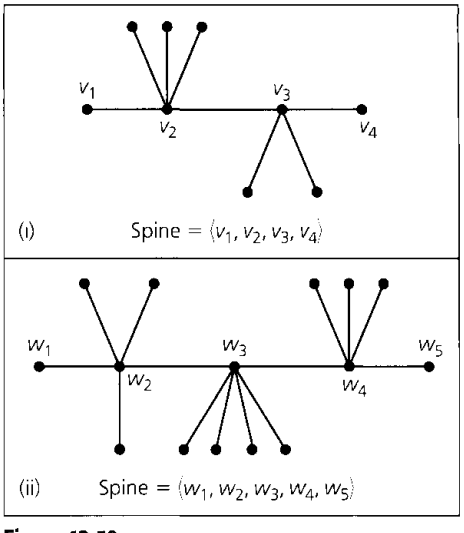


Figure 12.50

16. In part (i) of Fig. 12.51 we find a graceful labeling of the caterpillar shown in part (i) of Fig. 12.50. Find a graceful labeling for the caterpillars in part (ii) of Figs. 12.50 and 12.51.

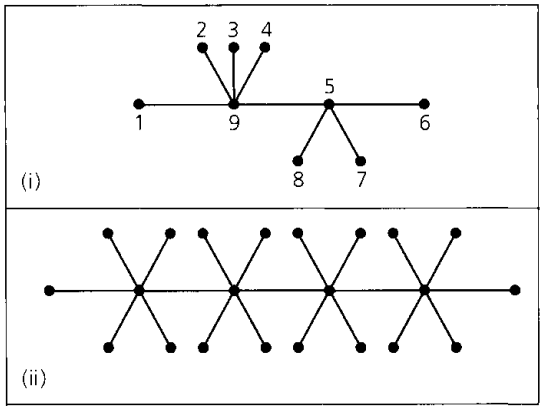


Figure 12.51

17. Develop an algorithm to gracefully label the vertices of a caterpillar with at least two edges.

18. Consider the caterpillar in part (i) of Fig. 12.50. If we label each edge of the spine with a 1 and each of the other edges with a 0, the caterpillar can be represented by a binary string. Here that binary string is 10001001 where the first 1 is for the first (left-most) edge of the spine, the next three 0's are for the (nonspine) edges at v_2, the second 1 is for edge $\{v_2, v_3\}$, the two 0's are for the (nonspine) leaves at v_3, and the final 1 accounts for the third (right-most) edge of the spine.

We also note that the reversal of the binary string 10001001 — namely, 10010001 — corresponds with a second caterpillar that is isomorphic to the one in part (i) of Fig. 12.50.

a) Find the binary strings for each of the caterpillars in part (ii) of Figs. 12.50 and 12.51.

b) Can a caterpillar have a binary string of all 1's?

c) Can the binary string for a caterpillar have only two 1's?

d) Draw all the nonisomorphic caterpillars on five vertices. For each caterpillar determine its binary string. How many of these binary strings are palindromes?

e) Answer the question posed in part (d) upon replacing "five" by "six."

f) For $n \geq 3$, prove that the number of nonisomorphic caterpillars on n vertices is $(1/2)(2^{n-3} + 2^{\lceil (n-3)/2 \rceil}) = 2^{n-4} + 2^{\lfloor (n-4)/2 \rfloor} = 2^{n-4} + 2^{\lfloor n/2 \rfloor - 2}$. (This was first established in 1973 by F. Harary and A. J. Schwenk.)

19. For $n \geq 0$, we want to count the number of ordered rooted trees on $n + 1$ vertices. The five trees in Fig. 12.52(a) cover the case for $n = 3$.

[*Note*: Although the two trees in Fig. 12.52(b) are distinct as binary rooted trees, as ordered rooted trees they are considered the same tree and each is accounted for by the fourth tree in Fig. 12.52(a).]

a) Performing a postorder traversal of each tree in Fig. 12.52(a), we traverse each edge twice — once going down and once coming back up. When we traverse an edge going down, we shall write "1" and when we traverse one coming back up, we shall write "−1." Hence the postorder traversal for the first tree in Fig. 12.52(a) generates the list 1, 1, 1, −1, −1, −1. The list 1, 1, −1, −1, 1, −1 arises for the second tree in part (a) of the figure. Find the corresponding lists for the other three trees in Fig. 12.52(a).

b) Determine the ordered rooted trees on five vertices that generate the lists: (i) 1, −1, 1, 1, −1, 1, −1, −1; (ii) 1, 1, −1, −1, 1, 1, −1, −1; and (iii) 1, −1, 1, −1, 1, 1, −1, −1. How many such trees are there on five vertices?

c) For $n \geq 0$, how many ordered rooted trees are there for $n + 1$ vertices?

20. For $n \geq 1$, let t_n count the number of spanning trees for the *fan* on $n + 1$ vertices. The fan for $n = 4$ is shown in Fig. 12.53.

a) Show that $t_{n+1} = t_n + \sum_{i=0}^{n} t_i$, where $n \geq 1$ and $t_0 = 1$.

b) For $n \geq 2$, show that $t_{n+1} = 3t_n - t_{n-1}$.

c) Solve the recurrence relation in part (b) and show that for $n \geq 1$, $t_n = F_{2n}$, the $2n$th Fibonacci number.

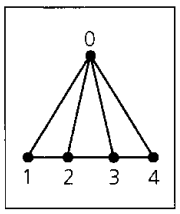


Figure 12.53

21. a) Consider the subgraph of G (in Fig. 12.54) induced by the vertices a, b, c, d. This graph is called a *kite*. How many nonidentical (though some may be isomorphic) spanning trees are there for this kite?

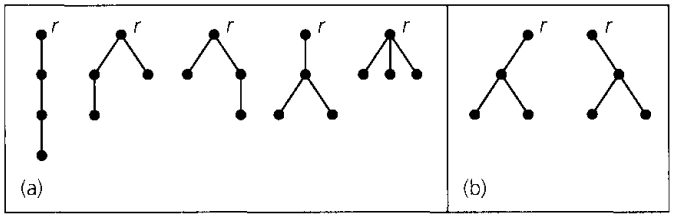


Figure 12.52

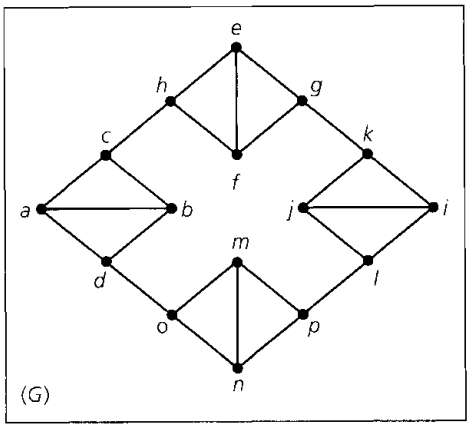


Figure 12.54

b) How many nonidentical (though some may be isomorphic) spanning trees of G do not contain the edge $\{c, h\}$?

c) How many nonidentical (though some may be isomorphic) spanning trees of G contain all four of the edges $\{c, h\}$, $\{g, k\}$, $\{l, p\}$, and $\{d, o\}$?

d) How many nonidentical (though some may be isomorphic) spanning trees exist for G?

e) We generalize the graph G as follows. For $n \geq 2$, start with a cycle on the $2n$ vertices $v_1, v_2, \ldots, v_{2n-1}, v_{2n}$. Replace each of the n edges $\{v_1, v_2\}$, $\{v_3, v_4\}, \ldots,$ $\{v_{2n-1}, v_{2n}\}$ with a (labeled) kite so that the resulting graph is 3-regular. (The case for $n = 4$ appears in Fig. 12.54.) How many nonidentical (though some may be isomorphic) spanning trees are there for this graph?