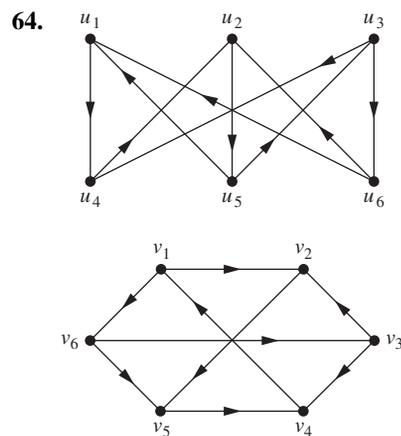
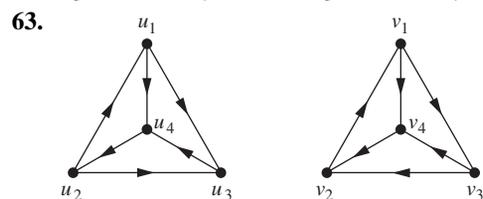
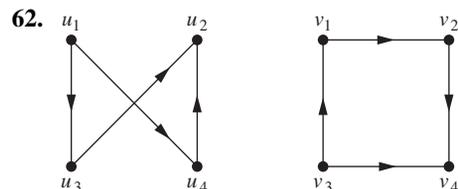
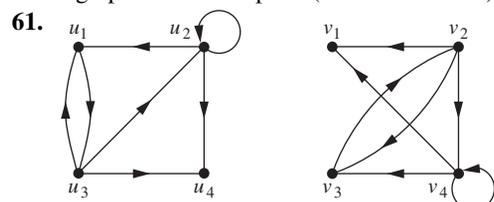


In Exercises 61–64 determine whether the given pair of directed graphs are isomorphic. (See Exercise 60.)



65. Show that if G and H are isomorphic directed graphs, then the converses of G and H (defined in the preamble of Exercise 67 of Section 10.2) are also isomorphic.

66. Show that the property that a graph is bipartite is an isomorphic invariant.

67. Find a pair of nonisomorphic graphs with the same degree sequence (defined in the preamble to Exercise 36 in Section 10.2) such that one graph is bipartite, but the other graph is not bipartite.

*68. How many nonisomorphic directed simple graphs are there with n vertices, when n is

- a) 2? b) 3? c) 4?

*69. What is the product of the incidence matrix and its transpose for an undirected graph?

*70. How much storage is needed to represent a simple graph with n vertices and m edges using

- a) adjacency lists?
- b) an adjacency matrix?
- c) an incidence matrix?

A **devil's pair** for a purported isomorphism test is a pair of nonisomorphic graphs that the test fails to show that they are not isomorphic.

71. Find a devil's pair for the test that checks the degree sequence (defined in the preamble to Exercise 36 in Section 10.2) in two graphs to make sure they agree.

72. Suppose that the function f from V_1 to V_2 is an isomorphism of the graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$. Show that it is possible to verify this fact in time polynomial in terms of the number of vertices of the graph, in terms of the number of comparisons needed.

10.4 Connectivity

Introduction

Many problems can be modeled with paths formed by traveling along the edges of graphs. For instance, the problem of determining whether a message can be sent between two computers using intermediate links can be studied with a graph model. Problems of efficiently planning routes for mail delivery, garbage pickup, diagnostics in computer networks, and so on can be solved using models that involve paths in graphs.

Paths

Informally, a **path** is a sequence of edges that begins at a vertex of a graph and travels from vertex to vertex along edges of the graph. As the path travels along its edges, it visits the vertices along this path, that is, the endpoints of these edges.

A formal definition of paths and related terminology is given in Definition 1.

DEFINITION 1

Let n be a nonnegative integer and G an undirected graph. A *path* of length n from u to v in G is a sequence of n edges e_1, \dots, e_n of G for which there exists a sequence $x_0 = u, x_1, \dots, x_{n-1}, x_n = v$ of vertices such that e_i has, for $i = 1, \dots, n$, the endpoints x_{i-1} and x_i . When the graph is simple, we denote this path by its vertex sequence x_0, x_1, \dots, x_n (because listing these vertices uniquely determines the path). The path is a *circuit* if it begins and ends at the same vertex, that is, if $u = v$, and has length greater than zero. The path or circuit is said to *pass through* the vertices x_1, x_2, \dots, x_{n-1} or *traverse* the edges e_1, e_2, \dots, e_n . A path or circuit is *simple* if it does not contain the same edge more than once.

When it is not necessary to distinguish between multiple edges, we will denote a path e_1, e_2, \dots, e_n , where e_i is associated with $\{x_{i-1}, x_i\}$ for $i = 1, 2, \dots, n$ by its vertex sequence x_0, x_1, \dots, x_n . This notation identifies a path only as far as which vertices it passes through. Consequently, it does not specify a unique path when there is more than one path that passes through this sequence of vertices, which will happen if and only if there are multiple edges between some successive vertices in the list. Note that a path of length zero consists of a single vertex.

Remark: There is considerable variation of terminology concerning the concepts defined in Definition 1. For instance, in some books, the term **walk** is used instead of *path*, where a walk is defined to be an alternating sequence of vertices and edges of a graph, $v_0, e_1, v_1, e_2, \dots, v_{n-1}, e_n, v_n$, where v_{i-1} and v_i are the endpoints of e_i for $i = 1, 2, \dots, n$. When this terminology is used, **closed walk** is used instead of *circuit* to indicate a walk that begins and ends at the same vertex, and **trail** is used to denote a walk that has no repeated edge (replacing the term *simple path*). When this terminology is used, the terminology **path** is often used for a trail with no repeated vertices, conflicting with the terminology in Definition 1. Because of this variation in terminology, you will need to make sure which set of definitions are used in a particular book or article when you read about traversing edges of a graph. The text [GrYe06] is a good reference for the alternative terminology described in this remark.

EXAMPLE 1

In the simple graph shown in Figure 1, a, d, c, f, e is a simple path of length 4, because $\{a, d\}$, $\{d, c\}$, $\{c, f\}$, and $\{f, e\}$ are all edges. However, d, e, c, a is not a path, because $\{e, c\}$ is not an edge. Note that b, c, f, e, b is a circuit of length 4 because $\{b, c\}$, $\{c, f\}$, $\{f, e\}$, and $\{e, b\}$ are edges, and this path begins and ends at b . The path a, b, e, d, a, b , which is of length 5, is not simple because it contains the edge $\{a, b\}$ twice. ▶

Paths and circuits in directed graphs were introduced in Chapter 9. We now provide more general definitions.

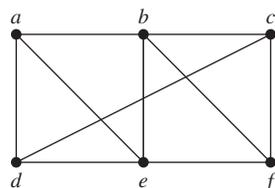


FIGURE 1 A Simple Graph.

DEFINITION 2

Let n be a nonnegative integer and G a directed graph. A *path* of length n from u to v in G is a sequence of edges e_1, e_2, \dots, e_n of G such that e_1 is associated with (x_0, x_1) , e_2 is associated with (x_1, x_2) , and so on, with e_n associated with (x_{n-1}, x_n) , where $x_0 = u$ and $x_n = v$. When there are no multiple edges in the directed graph, this path is denoted by its vertex sequence $x_0, x_1, x_2, \dots, x_n$. A path of length greater than zero that begins and ends at the same vertex is called a *circuit* or *cycle*. A path or circuit is called *simple* if it does not contain the same edge more than once.

Remark: Terminology other than that given in Definition 2 is often used for the concepts defined there. In particular, the alternative terminology that uses *walk*, *closed walk*, *trail*, and *path* (described in the remarks following Definition 1) may be used for directed graphs. See [GrYe05] for details.

Note that the terminal vertex of an edge in a path is the initial vertex of the next edge in the path. When it is not necessary to distinguish between multiple edges, we will denote a path e_1, e_2, \dots, e_n , where e_i is associated with (x_{i-1}, x_i) for $i = 1, 2, \dots, n$, by its vertex sequence x_0, x_1, \dots, x_n . The notation identifies a path only as far as which the vertices it passes through. There may be more than one path that passes through this sequence of vertices, which will happen if and only if there are multiple edges between two successive vertices in the list.

Paths represent useful information in many graph models, as Examples 2–4 demonstrate.

EXAMPLE 2

Paths in Acquaintanceship Graphs In an acquaintanceship graph there is a path between two people if there is a chain of people linking these people, where two people adjacent in the chain know one another. For example, in Figure 6 in Section 10.1, there is a chain of six people linking Kamini and Ching. Many social scientists have conjectured that almost every pair of people in the world are linked by a small chain of people, perhaps containing just five or fewer people. This would mean that almost every pair of vertices in the acquaintanceship graph containing all people in the world is linked by a path of length not exceeding four. The play *Six Degrees of Separation* by John Guare is based on this notion. ◀

EXAMPLE 3

Paths in Collaboration Graphs In a collaboration graph, two people a and b are connected by a path when there is a sequence of people starting with a and ending with b such that the endpoints of each edge in the path are people who have collaborated. We will consider two particular collaboration graphs here. First, in the academic collaboration graph of people who have written papers in mathematics, the **Erdős number** of a person m (defined in terms of relations in Supplementary Exercise 14 in Chapter 9) is the length of the shortest path between m and the extremely prolific mathematician Paul Erdős (who died in 1996). That is, the Erdős number of a mathematician is the length of the shortest chain of mathematicians that begins with Paul Erdős and ends with this mathematician, where each adjacent pair of mathematicians have written a joint paper. The number of mathematicians with each Erdős number as of early 2006, according to the Erdős Number Project, is shown in Table 1.

In the Hollywood graph (see Example 3 in Section 10.1) two actors a and b are linked when there is a chain of actors linking a and b , where every two actors adjacent in the chain have acted in the same movie. In the Hollywood graph, the **Bacon number** of an actor c is defined to be the length of the shortest path connecting c and the well-known actor Kevin Bacon. As new movies are made, including new ones with Kevin Bacon, the Bacon number of actors can change. In Table 2 we show the number of actors with each Bacon number as of early 2011 using data from the Oracle of Bacon website. The origins of the Bacon number of an actor dates back to the early 1990s, when Kevin Bacon remarked that he had worked with everyone in Hollywood or someone who worked with them. This led some people to invent a party

Replace Kevin Bacon by your own favorite actor to invent a new party game

TABLE 1 The Number of Mathematicians with a Given Erdős Number (as of early 2006).

<i>Erdős Number</i>	<i>Number of People</i>
0	1
1	504
2	6,593
3	33,605
4	83,642
5	87,760
6	40,014
7	11,591
8	3,146
9	819
10	244
11	68
12	23
13	5

TABLE 2 The Number of Actors with a Given Bacon Number (as of early 2011).

<i>Bacon Number</i>	<i>Number of People</i>
0	1
1	2,367
2	242,407
3	785,389
4	200,602
5	14,048
6	1,277
7	114
8	16

game where participants were challenged to find a sequence of movies leading from each actor named to Kevin Bacon. We can find a number similar to a Bacon number using any actor as the center of the acting universe. ◀

Connectedness in Undirected Graphs

When does a computer network have the property that every pair of computers can share information, if messages can be sent through one or more intermediate computers? When a graph is used to represent this computer network, where vertices represent the computers and edges represent the communication links, this question becomes: When is there always a path between two vertices in the graph?

DEFINITION 3

An undirected graph is called *connected* if there is a path between every pair of distinct vertices of the graph. An undirected graph that is not *connected* is called *disconnected*. We say that we *disconnect* a graph when we remove vertices or edges, or both, to produce a disconnected subgraph.

Thus, any two computers in the network can communicate if and only if the graph of this network is connected.

EXAMPLE 4

The graph G_1 in Figure 2 is connected, because for every pair of distinct vertices there is a path between them (the reader should verify this). However, the graph G_2 in Figure 2 is not connected. For instance, there is no path in G_2 between vertices a and d . ◀

We will need the following theorem in Chapter 11.

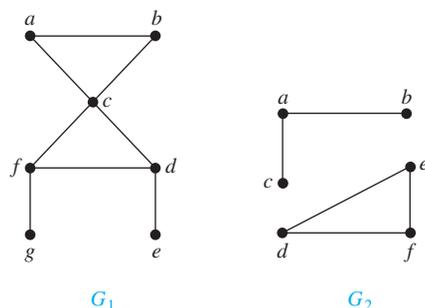


FIGURE 2 The Graphs G_1 and G_2 .

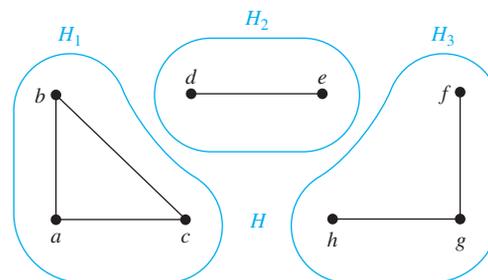


FIGURE 3 The Graph H and Its Connected Components H_1 , H_2 , and H_3 .

THEOREM 1

There is a simple path between every pair of distinct vertices of a connected undirected graph.

Proof: Let u and v be two distinct vertices of the connected undirected graph $G = (V, E)$. Because G is connected, there is at least one path between u and v . Let x_0, x_1, \dots, x_n , where $x_0 = u$ and $x_n = v$, be the vertex sequence of a path of least length. This path of least length is simple. To see this, suppose it is not simple. Then $x_i = x_j$ for some i and j with $0 \leq i < j$. This means that there is a path from u to v of shorter length with vertex sequence $x_0, x_1, \dots, x_{i-1}, x_j, \dots, x_n$ obtained by deleting the edges corresponding to the vertex sequence x_i, \dots, x_{j-1} . ◀

CONNECTED COMPONENTS A **connected component** of a graph G is a connected subgraph of G that is not a proper subgraph of another connected subgraph of G . That is, a connected component of a graph G is a maximal connected subgraph of G . A graph G that is not connected has two or more connected components that are disjoint and have G as their union.



EXAMPLE 5

What are the connected components of the graph H shown in Figure 3?

Solution: The graph H is the union of three disjoint connected subgraphs H_1 , H_2 , and H_3 , shown in Figure 3. These three subgraphs are the connected components of H . ◀

EXAMPLE 6

Connected Components of Call Graphs Two vertices x and y are in the same component of a telephone call graph (see Example 4 in Section 10.1) when there is a sequence of telephone calls beginning at x and ending at y . When a call graph for telephone calls made during a particular day in the AT&T network was analyzed, this graph was found to have 53,767,087 vertices, more than 170 million edges, and more than 3.7 million connected components. Most of these components were small; approximately three-fourths consisted of two vertices representing pairs of telephone numbers that called only each other. This graph has one huge connected component with 44,989,297 vertices comprising more than 80% of the total. Furthermore, every vertex in this component can be linked to any other vertex by a chain of no more than 20 calls. ◀



How Connected is a Graph?

Suppose that a graph represents a computer network. Knowing that this graph is connected tells us that any two computers on the network can communicate. However, we would also like to understand how reliable this network is. For instance, will it still be possible for all computers to communicate after a router or a communications link fails? To answer this and similar questions, we now develop some new concepts.



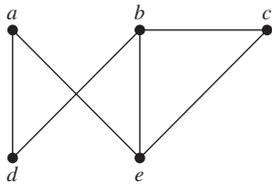
there are exactly eight paths of length four from a to d . By inspection of the graph, we see that $a, b, a, b, d; a, b, a, c, d; a, b, d, b, d; a, b, d, c, d; a, c, a, b, d; a, c, a, c, d; a, c, d, b, d;$ and a, c, d, c, d are the eight paths of length four from a to d .

Theorem 2 can be used to find the length of the shortest path between two vertices of a graph (see Exercise 56), and it can also be used to determine whether a graph is connected (see Exercises 61 and 62).

Exercises

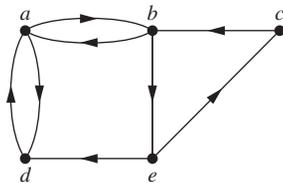
- Does each of these lists of vertices form a path in the following graph? Which paths are simple? Which are circuits? What are the lengths of those that are paths?

a) a, e, b, c, b	b) a, e, a, d, b, c, a
c) e, b, a, d, b, e	d) c, b, d, a, e, c

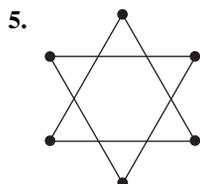
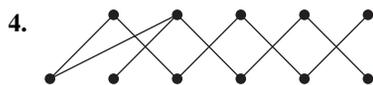


- Does each of these lists of vertices form a path in the following graph? Which paths are simple? Which are circuits? What are the lengths of those that are paths?

a) a, b, e, c, b	b) a, d, a, d, a
c) a, d, b, e, a	d) a, b, e, c, b, d, a



In Exercises 3–5 determine whether the given graph is connected.



- How many connected components does each of the graphs in Exercises 3–5 have? For each graph find each of its connected components.
- What do the connected components of acquaintanceship graphs represent?
- What do the connected components of a collaboration graph represent?
- Explain why in the collaboration graph of mathematicians (see Example 3 in Section 10.1) a vertex representing a mathematician is in the same connected component as the vertex representing Paul Erdős if and only if that mathematician has a finite Erdős number.
- In the Hollywood graph (see Example 3 in Section 10.1), when is the vertex representing an actor in the same connected component as the vertex representing Kevin Bacon?
- Determine whether each of these graphs is strongly connected and if not, whether it is weakly connected.

