CHAPTER 5

Networks and Graphs

In the last chapter we borrowed from language used to describe mathematical relations to describe relations among people. For example, human relationships, like authority in organizations, can be transitive in the same sense as the "greater than" relation between numbers. In this chapter we introduce another branch of mathematics that has proven to be very useful in describing relations between people: *graph theory*. Graph theory, contrary to what you might think, has nothing to do with statistical plots of data. Graph theory is all about points, called *vertices* or *nodes* (we will use these terms interchangeably), and the lines connecting them, called *edges* or *arcs*. Graphs are diagrams of these vertices and lines. Graphs are diagrams of networks because the vertices are the objects in the network (people, countries, computers, etc.) and the edges are relationships. For example, consider the graph in Figure 5.1, known as a star.

This is a graph with five vertices and four edges. Vertices connected by an edge are said to be adjacent to one another. The number of vertices adjacent to the given vertex is called its *degree*. In this star graph one vertex has degree four and four vertices have degree one.

We can add some flesh to this example if the vertices are people and the edges are some symmetric relation among these people. For example, the relation might be mutual friendship. Figure 5.2 shows a situation in which one person has four friends who are not friends with each other. The degree here has a simple interpretation: it is a measure of each person's popularity. This group is all about Eve—she's the most popular.

This is not a very tight group of friends; most of the pairs are not friends of each other. There are only four friendships among these five people. A little thought will show that there are six missing relations in this group, those among Annie, Drew, Barbara, and Charlotte. A simple measure of the *cohesion* of a group is the number of relations to the total possible relations—4/10 or .40 in this group. This measure of density varies from .00, when no one is connected to anyone else, to 1.00, when all possible pairs are connected. If there are N people in a group, the number of possible relations is $N(N-1)/2$. In this case, $N = 5$ and $5 \times 4/2 = 10$.

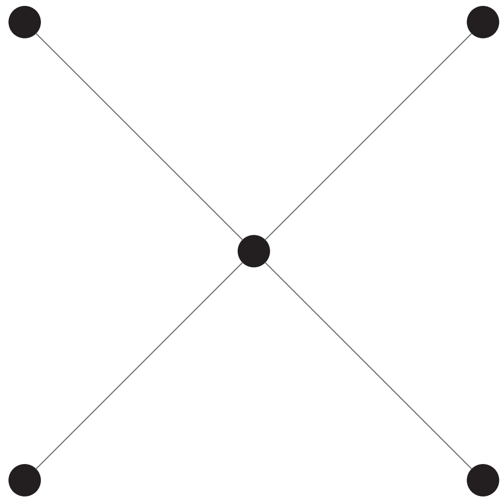

Figure 5.1. A graph with five nodes and four edges.

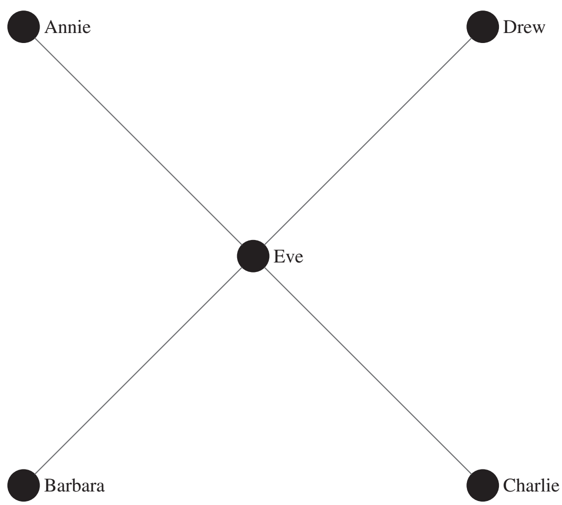

Figure 5.2. A network of five persons and four relationships.

$$\text{Number of possible edges in a graph with N vertices} = \frac{N(N-1)}{2} \tag{5.1}$$

$$\text{Density} = \frac{\text{Number of Edges}}{\frac{N(N-1)}{2}} \tag{5.2}$$

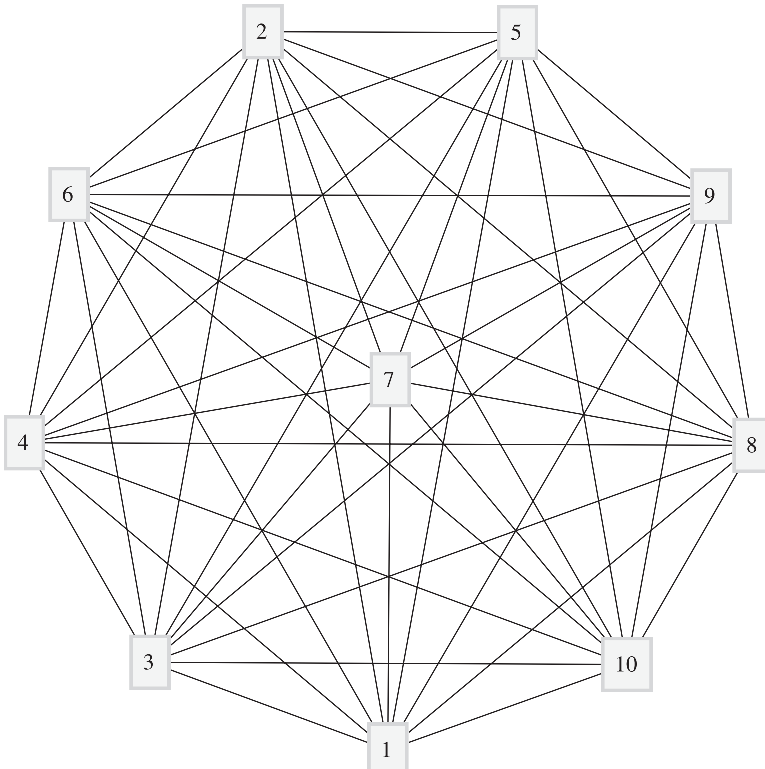

Figure 5.3. A complete network.

We will explore the idea of density further in the chapter on small-world networks.

Notice that although Barbara and Charlotte are not friends, they share a friend. This is important because often in networks things can be transmitted. For example, if Charlotte hears a rumor and friends share information, then Eve will learn of the rumor also. Charlotte may even tell Eve secrets she has learned she has promised to reveal to no one else. If Charlotte catches a cold, Barbara may be exposed to it also through Eve indirectly. A *walk* is a sequence of connected edges that indirectly connects two nodes. Barbara and Charlotte are connected by many different walks: Barbara-Eve-Charlotte, Barbara-Eve-Drew-Eve-Charlotte, Barbara-Eve-Annie-Eve-Drew-Eve-Charlotte, and so forth. Some walks traverse the same edges or the same nodes more than once. Walks that do not repeat either edges or nodes are called *paths*. There is just one path connecting Barbara and Charlotte, although in other more complex networks there may me more than one path connecting two nodes.

A star is one particularly simple kind of network. Another simple type is a complete network, in which all pairs are connected. Figure 5.3 gives a

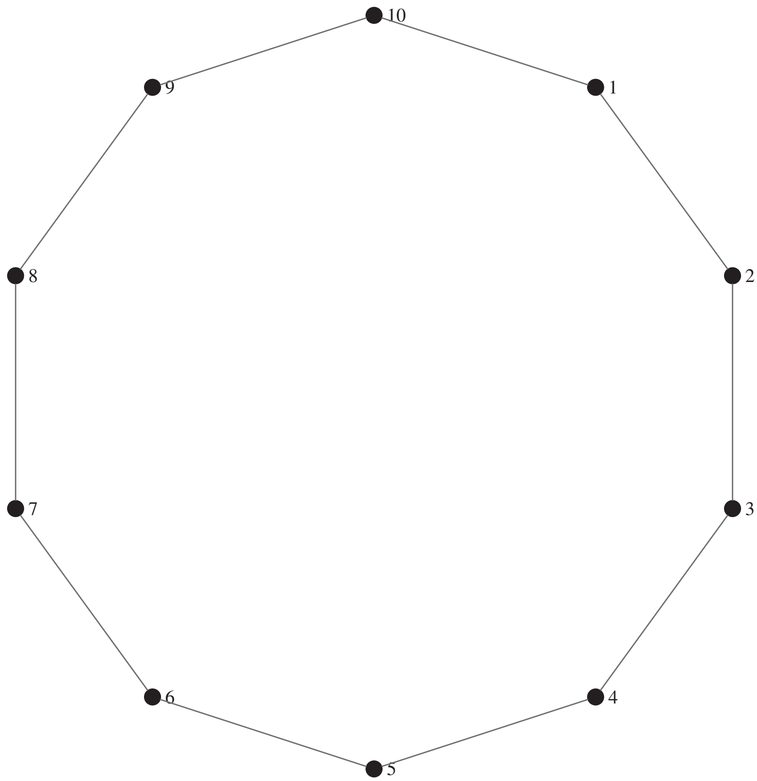

Figure 5.4. A cycle network.

complete network with 10 nodes. Note that even though Node 7 lies in the middle of the graph, the positioning is arbitrary. In a complete network, all nodes are essentially the same in the respect that they have a connection to every other node.

Another simple kind of network is a cycle, a network in which the nodes form a line with the first and last nodes connected to each other. Figure 5.4 gives a cycle network with 10 nodes.

Notice that in this network there are two paths connecting every pair of nodes, but for some pairs the two paths are of unequal length. For example, Nodes 1 and 3 are connected by a path with two edges, 1-2-3, and a path with eight edges, 1-10-9-8-7-6-5-4-3. The length of the shortest path between two nodes is the distance between the nodes, and the shortest paths connecting two nodes are called *geodesics*. Note that some pairs of vertices are connected by two different geodesics.

A grid is a network that can be arranged into a two-dimensional square or rectangle block-like structure. Figure 5.5 gives a grid network with 25 nodes.

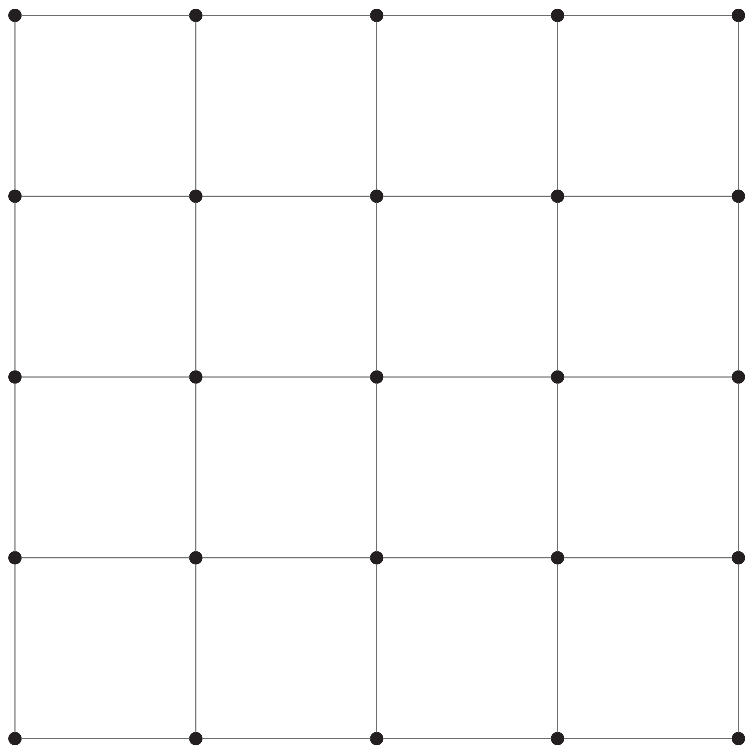

Figure 5.5. A grid network.

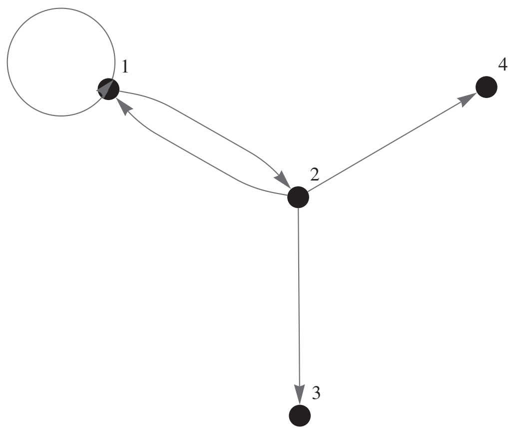

Figure 5.6. A graph that is not simple.

All the networks we will be looking at in this book are *simple*. The word "simple" has a technical meaning. There are no loops (edges connecting a node to itself), and there is at most one edge connecting any pair of vertices. Examine the network in Figure 5.6. There are two reasons why it is not a simple network.

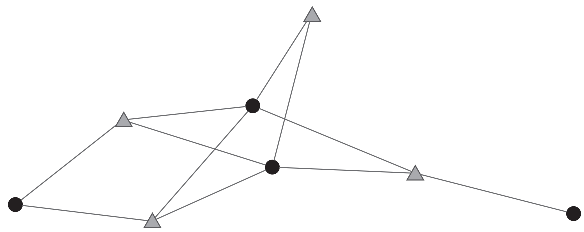

Figure 5.7. A bimodal (bipartite) network.

Some graphs involve two types of complementary nodes in which all connections are between nodes of different types, with none being between nodes of the same type. These are called *bipartite* graphs. For example, many sociologists study networks of directorships of corporations, where there are connections between directors and the corporations they direct. Figure 5.7 shows a hypothetical bipartite dating network between young men and women.

Finding a bipartite network can be an important finding. Suppose you were a newspaper reporter doing a study of expensive Washington restaurants and you observed who ate with whom. You find that almost every pair involved a congressman and a lobbyist. The lobbyist buys an expensive dinner for the congressman in order to influence her vote. You might find that two lobbyists or two congressmen never dine together in these costly restaurants. What's the point! This is the kind of finding you would certainly include in our newspaper article.

It may happen that there is no path between some pairs of vertices; you can't reach one from the other even through indirect paths. This is important because it means that nothing can be transmitted from one vertex to the other. It means, for example, that a contagious disease cannot be transmitted from one indirectly to the other or that a rumor started by one will not reach the other. Sets of vertices that are mutually reachable form a *component* of a graph.

The networks (graphs) we've been looking at thus far have edges, which are bi-directional: A are B are friends of one another, or talk to one another. But some relations can be one directional, and they are represented by digraphs with arcs, not edges. Suppose, for example, that everyone talks about their personal problems with Eve, but she does not reveal any of hers to them. Then the network could be described by the diagram in Figure 5.8, which has four arcs and no edges.

Authority relations in an organizational can also be unsymmetrical. In Figure 5.9 there are three levels to a hierarchy. Positions 1 and 3 both have subordinates. Figures 5.8 is also an example of a trees: it is a network without any cycles. We will talk more about trees in later chapters.

Note that in digraphs the connections of any node are of two types: those that flow in and those that flow out. In Figure 5.8 Eve has four people

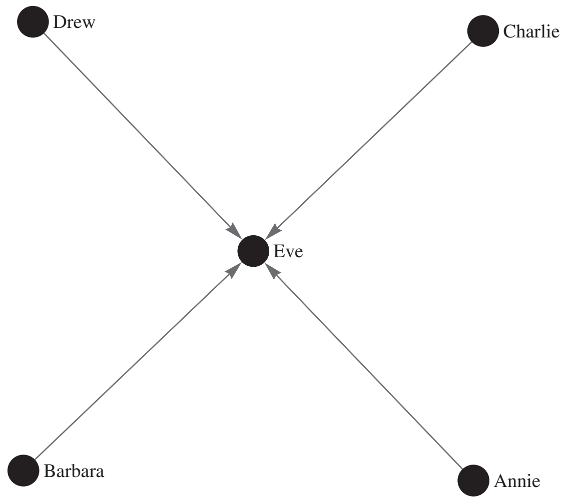

Figure 5.8. A digraph.

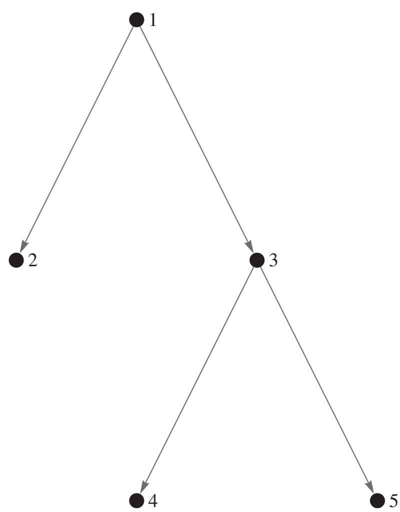

Figure 5.9. A digraph that is a tree.

who talk to her and no one she confides in. In Figure 5.9 Position 3 has one superior and two subordinates. In digraphs we distinguish between in-degree and out-degree. Eve has an *in-degree* of four and an *out-degree* of zero.

EXERCISES

1. Draw all the distinct tree graphs with five vertices. Do they have different densities? Do they have different distributions of degrees? Do some of them have more inequality of degree than others?

Table 5.1.
Graph with 50 nodes.

Average degree	Mostly one component? (yes/no)
0.50	
0.75	
1.00	
1.25	
1.50	
1.75	
2.00	

Table 5.2.
Graph with 100 nodes.

Average degree	Mostly one component? (yes/no)
0.50	
0.75	
1.00	
1.25	
1.50	
1.75	
2.00	

2. Draw a network with five nodes in which the distances between all pairs of nodes are equal.
3. In what size network is the maximum number of edges?
4. What is the maximum number of arcs in a digraph with five vertices? In a digraph with N vertices?
5. What is the maximum distance between two nodes in a network with 10 nodes?
6. Why is the network in Figure 6 not “simple”?
7. Show that the possible components of a graph are mutually exclusive, that they never overlap, and that a vertex cannot belong to two different components.
8. The demonstration Random Graphs creates random networks in which you can control the number of nodes and the average degree. The “random choice” slider will enable you to generate multiple random networks. Experiment with graphs with 50 nodes and an average degree of .50. Is there typically just one component or more than one? Gradually increase the average degree. At what point do most off the random graphs consist of just one component? Now do the same thing for graphs of size 100. Use Tables 5.1 and 5.2.