Part II

Mathematical Representations of Social Networks

3
Notation for Social Network Data

Social network data consist of measurements on a variety of relations for one or more sets of actors. In a network data set we may also have recorded information on attributes of the actors. We will need notation for the set of actors, the relations themselves, and the actor attributes so that we can refer to important network concepts in a unified manner.

In this chapter, we introduce notation and illustrate with examples. We start by defining notation for a single, dichotomous relation. We then move to more complicated network data sets involving more than one set of actors and/or more than one relation. We will need a notational system flexible enough to handle the wide range of network data sets that are encountered in practice. We note that the only type of structural variable discussed in this chapter is relational. Chapter 8 presents notation and methodology for affiliational networks.

For the reader who already is familiar with social networks and the ways in which social network data can be denoted, or the reader who is only interested in specific techniques, we recommend a quick reading of the material in this chapter. Specifically, such readers can glance at Section 2 and the examples used in this chapter (perhaps skipping the material on multiple relations), and return to this chapter as needed.

There are many ways to describe social network data mathematically. We will introduce three different notational schemes. These schemes can each be adapted to represent a wide range of network data. However, for some forms of data and some types of network methods, one notation scheme may be preferred to the others, because of its appropriateness, clarity, or efficiency. The notations are:

- Graph theoretic
- Sociometric

- Algebraic

Each scheme will be described and illustrated in detail. We will show how these schemes overlap, and discuss when a specific scheme is more useful than the others. Graph theoretic notation is most useful for centrality and prestige methods, cohesive subgroup ideas, as well as dyadic and triadic methods. Sociometric notation is often used for the study of structural equivalence and blockmodels. Algebraic notation is most appropriate for role and positional analyses and relational algebras. We should note that there are other ways to denote social network data, some of which are used to study specific statistical models. Such schemes will be mentioned, when needed, in later chapters.

The graph theoretic notation scheme can be viewed as an elementary way to represent actors and relations. It is the basis of the many concepts of graph theory used since the late 1940's to study social networks. The notation provides a straightforward way to refer to actors and relations, and is completely consistent with the notation from the other three schemes. Mathematicians and statisticians such as Bock, Harary, Katz, and Luce were among the first to view networks as directed and undirected graphs (see Forsyth and Katz 1946; Katz 1947; Luce and Perry 1949; Bock and Husain 1950, 1952; Harary and Norman 1953). Graph theory texts such as Flament (1963) and Harary (1969) describe social network applications. We should also direct the reader to other texts on graph theory and social networks, such as Harary, Norman, and Cartwright (1965), and Hage and Harary (1983), that present graph theoretic notation for social network data. Mathematical sociology texts, such as Coleman (1964), Fararo (1973), and Leik and Meeker (1975), contain elementary discussions of the use of graph theory in social network analysis.

The second notation scheme, sociometric notation, is by far the most common in the social network literature. One presents the data for each relation in a two-way matrix, termed a *sociomatrix*, where the rows and columns refer to the actors making up the pairs. Sociomatrices began to be used more than fifty years ago after their introduction by Moreno (1934) in his pioneering research in sociometry (see also Moreno and Jennings 1938).

Most major computer software packages for social network data analyze network information presented in sociomatrices. Further, many methods are defined for sociomatrices. This notational scheme is probably the most useful for readers interested in the methods discussed in Parts III and IV of the book. Sociomatrices are *adjacency matrices* for *graphs*, and consequently, this second notational scheme is directly related to the first.

The third notational scheme, algebraic notation, is used to study multiple relations. This notation is useful for studying network role structures and relational algebras. Such analyses use algebraic techniques to compare and contrast measured relations, and derived compound relations. A compound relation is the composition or combination of two or more relations. For example, if we have measured two relations, "is a friend of" and "is an enemy of," for a set of people, then we might be interested in the composition of these two relations: "friends' enemies." The focus of such algebraic techniques is on the associations among the relations measured on pairs of actors, across the entire set of actors. This notation is designed for one-mode networks, and was first used by White (1963) and Boyd (1969).

We now turn to each of these notations, show how they are related, discuss when each is useful, and illustrate each with examples.

3.1 Graph Theoretic Notation

A network can be viewed in several ways. One of the most useful views is as a *graph*, consisting of *nodes* joined by *lines*. Chapter 4 discusses graph theory at length. Here, we introduce some simple graph theoretic notation, and show how this notation can be used to label the actors and relations in a network data set.

Suppose we have a set of actors. We will refer to this set as $\mathcal{N}$. The set $\mathcal{N}$ contains g actors in number, which we will denote by $\mathcal{N} = \{n_1, n_2, \ldots, n_g\}$. The symbol $\mathcal{N}$ is commonly used to stand for the set, since the graph theory literature frequently refers to this set as a collection of *nodes* of a graph. For example, consider a collection of $g = 6$ second-grade children: Allison, Drew, Eliot, Keith, Ross, and Sarah. We have $\mathcal{N}$ = {Allison, Drew, Eliot, Keith, Ross, Sarah}, a collection of six actors, so that we can refer to the children by their symbols: n_1 = Allison, n_2 = Drew, n_3 = Eliot, n_4 = Keith, n_5 = Ross, and n_6 = Sarah.

3.1.1 A Single Relation

We now assume that we have a single relation for the set of actors $\mathcal{N}$. That is, we record whether each actor in $\mathcal{N}$ relates to every other actor on this relation. To start, we will let the relation be dichotomous and directional. Thus, n_i either relates to n_j or does not. For now, we do not consider the strength of this interaction or how frequently n_i interacts with n_j.

Consider an ordered pair of actors, n_i and n_j. Either the first actor in the ordered pair relates to the second or it does not. Since the relation is directional, the pair of actors n_i and n_j is distinct from the pair n_j and n_i (that is, order matters). If a tie is present, then we say that the ordered pair is an element of a special collection of pairs, which we will refer to as $\mathscr{L}$. If an ordered pair is in $\mathscr{L}$, then the first actor in the pair relates to the second on the relation under consideration.

Note that there can be as many as $g(g-1)$ elements (the total number of ordered pairs in $\mathscr{L}$), and as few as 0.

If the ordered pair under consideration is $< n_i, n_j >$, and if there is a tie present, we will write $n_i \rightarrow n_j$. The elements, or ordered pairs, of relating actors in $\mathscr{L}$ will be denoted by the symbol l. Let us assume that there are L entries in $\mathscr{L}$, so that $\mathscr{L} = \{l_1, l_2, \ldots, l_L\}$. The elements in $\mathscr{L}$ can be represented graphically by drawing a line from the first actor in the element to the second. It is customary to refer to such a graph as a *directed graph*, since the lines have a direction. Directed lines are referred to as *arcs*. We use the symbol $\mathscr{L}$ to refer to the set of directed *Lines* and the symbol l to refer to the individual directed *lines* in the set. We will frequently refer to such ordered pairs of relating actors as *directed lines* or *arcs*.

Since a graph consists of a set of nodes $\mathscr{N}$, and a set of lines $\mathscr{L}$, it can be described mathematically by the two sets, $(\mathscr{N}, \mathscr{L})$. We will use the symbol $\mathscr{G}$ to denote a graph. It is important to note that for the graph theoretic notation scheme, these two sets (a set of actors, and a set of ordered pairs of actors, or arcs) suffice for mathematical descriptions of the crucial components in a network on which a single, dichotomous relation is measured.

On some relations, an individual actor does not usually relate to itself. When studying such relations, one does not consider *self-choices*.

There are relations that are *nondirectional*; that is, we cannot distinguish between the line from n_i to n_j and the line from n_j to n_i. For example, we may consider a set of actors, and record whether they "live near each other." Clearly, this is a nondirectional relation — if n_i lives near n_j, then n_j lives near n_i. There is only one measurement to be made for each pair, rather than two as with a directional relation. The two ordered pairs have identical relational interactions. The set $\mathscr{L}$ now contains at most $g(g-1)/2$ pairs. The order of the pair of actors in these relating pairs no longer matters, since both actors relate to each other in the same way.

Return to our example, and suppose that the single, dichotomous directional relation is "friendship," so that we consider whether each child views every other child as a friend. Suppose further that eight of the possible thirty ordered pairs are friendships (that is, eight of the thirty possible arcs are present) and that the other twenty-two are not friendships (or that there are twenty-two arcs absent). Let these $L = 8$ pairs be <Allison, Drew>, <Allison, Ross>, <Drew, Sarah>, <Drew, Eliot>, <Eliot, Drew>, <Keith, Ross>, <Ross, Sarah>, and <Sarah, Drew>. Thus, for the elements of $\mathscr{L}$, l_1 = <Allison, Drew>, l_2 = <Allison, Ross>, ... , and l_8 = <Sarah, Drew>. The data tell us that Allison views Drew as a friend, Allison also views Ross as a friend, Drew states that Sarah is his friend, and so forth. It is also interesting to note that this friendship is not reciprocal; that is, if n_i states that n_j is his friend (or $n_i \rightarrow n_j$), it is possible that this sentiment is not returned — n_j may not "choose" n_i as a friend (or $n_j \nrightarrow n_i$).

A graph can be presented as a diagram in which nodes are represented as points in a two-dimensional space and arcs are represented by directed arrows between points. Thus, these six children can be represented as points in a two-dimensional space. It is important to note that the actual location of points in this two-dimensional space is irrelevant. We can take these points, and draw in the eight arcs representing these eight ordered pairs of children who are friends. This directed graph or *sociogram* is shown in Figure 3.1.

3.1.2 ○Multiple Relations

We may have more than one relation in a social network data set. Graph theoretic notation can be generalized to multirelational networks, which could include both directional and nondirectional relations. For example, we may study whether the corporations in a metropolitan area do business with each other — does n_i sell to n_j, for example — and whether they interlock through their boards of directors — does an officer of corporation n_i sit on the board of directors of corporation n_j? Given the notation presented for the case of a single dichotomous relation, it is easy to generalize it to multiple relations.

Suppose that we are interested in more than one relation defined on pairs of actors taken from $\mathscr{N}$. Let R be the number of relations.

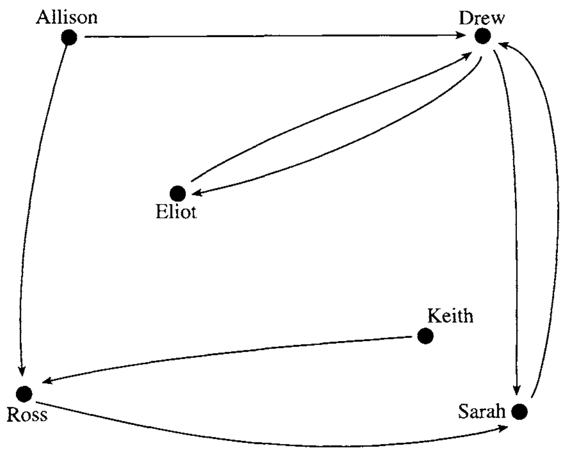


Fig. 3.1. The six actors and the directed lines between them — a sociogram

Each of these relations can be represented as a graph or directed graph; hence, each has associated with it a set of lines or arcs, specifying which (directed) lines are present in the (directed) graph for the relation (or, which (ordered) pairs are "relating"). Thus, each relation has a corresponding set of arcs, $\mathscr{L}_r$, which contains L_r ordered pairs of actors as elements. Here, the subscript r ranges from 1 to R, the total number of relations.

Each of these R sets defines a directed graph on the nodes in $\mathscr{N}$. These directed graphs can be viewed in one or more figures. So, each relation is defined on the same set of nodes, but each has a different set of arcs. Thus, we can quantify the rth relation by $(\mathscr{N}, \mathscr{L}_r)$, for $r = 1, 2, \ldots, R$.

For example, return to our second-graders, and now consider $R = 3$ relations: 1) who chooses whom as a friend, measured at the beginning of the school year; 2) who chooses whom as a friend, measured at the end of the school year; and 3) who lives near whom. The first two relations are directional, while the last is nondirectional. Suppose that $L_1 = 8$ ordered pairs of actors, $L_2 = 11$, and $L_3 = 12$. Below, we list these three sets.

Relation 1 Friendship at Beginning	Relation 2 Friendship at End	Relation 3 Lives Near
<Allison, Drew>	<Allison, Drew>	(Allison, Ross)
<Allison, Ross>	<Allison, Ross>	(Allison, Sarah)
<Drew, Sarah>	<Drew, Sarah>	(Drew, Eliot)
<Drew, Eliot>	<Drew, Eliot>	(Keith, Ross)
<Eliot, Drew>	<Drew, Ross>	(Keith, Sarah)
<Keith, Ross>	<Eliot, Ross>	(Ross, Sarah)
<Ross, Sarah>	<Keith, Drew>	
<Sarah, Drew>	<Keith, Ross>	
	<Ross, Keith>	
	<Ross, Sarah>	
	<Sarah, Drew>	

For a nondirectional relation, such as "lives near," measurements are made on unordered rather than ordered pairs. Clearly, when one actor relates to a second, the second relates to the first; therefore, since Allison lives near Ross, Ross lives near Allison. When listing the pairs of relating actors (or arcs) for a nondirectional relation, each pair can be listed no more than once. We use $(\bullet, \bullet)$ to denote pairs of actors for whom a tie is present on a nondirectional relation, and use $< \bullet, \bullet >$ to denote ties on a directional relation.

Examining such lists can be difficult. An alternative way to present the three sets $\mathscr{L}_1$, $\mathscr{L}_2$, and $\mathscr{L}_3$ is graphically. We can place the arcs for directed graphs or lines for undirected graphs on three figures (one for each relation), or on a single figure containing points representing the six actors and arcs or lines for all relations, simultaneously. We use different types of lines in Figure 3.2 for the different relations: solid, for relation 1 (friend at beginning); dashed, for relation 2 (friend at end); and dotted, for relation 3 (lives near). Since friendship is a directional relation, there are arrowheads indicating the directionality of an arc. Since "lives near" is nondirected, there are no arrowheads on these lines. This figure is an example of a *multivariate directed graph*; such graphs are described in more detail in Chapter 4.

3.1.3 Summary

To review, we have assumed that there is just one set of actors. This assumption will be relaxed in a later section of this chapter. In this simple

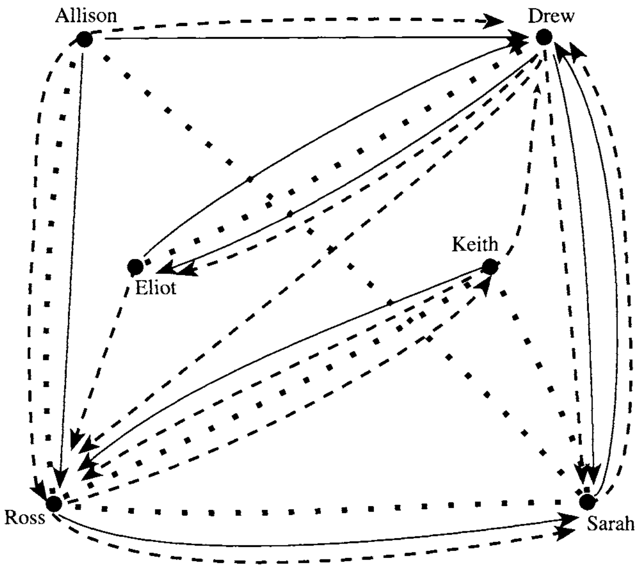


Fig. 3.2. The six actors and the three sets of directed lines — a multivariate directed graph

situation, there is just a single kind of pair of actors, those with both actors in the single set $\mathcal{N}$. The number of actors in $\mathcal{N}$ is g. Assuming that we have R relations, we have a set of arcs associated with each relation, $\mathscr{L}_1, \mathscr{L}_2, \ldots, \mathscr{L}_R$. Each set of arcs can have as many as $g(g-1)$ entries in these sets. The entries in each set are exactly those ordered pairs for which the first actor relates to the second actor on the relation in question. Thus, one needs to specify the set $\mathcal{N}$ and the R sets of arcs to describe completely the network data set.

We should mention that this notation scheme does not extend well to valued relations. Graph theory is not well designed for data sets that record the strength or frequency of the interaction for a pair of actors. One can use special graphs, such as signed graphs and valued graphs (see Chapter 4), to represent valued relations, but many of the more elegant results from graph theory do not apply to this extension. However, sociometric notation is general enough to handle valued relations.

3.2 Sociometric Notation

Sociometry is the study of positive and negative affective relations, such as liking/disliking and friends/enemies, among a set of people. A social network data set consisting of people and measured affective relations between people is often referred to as *sociometric*.

Relational data are often presented in two-way matrices termed *sociomatrices*. The two dimensions of a sociomatrix are indexed by the sending actors (the rows) and the receiving actors (the columns). Thus, if we have a one-mode network, the sociomatrix will be square.

A sociomatrix for a dichotomous relation is exactly the *adjacency matrix* for the graph (or *sociogram*) quantifying the ties between the actors for the relation in question. So, this notation can be viewed as complementary to graph theoretic notation described in the previous section. In these next paragraphs, we describe the history of sociomatrices and sociograms. We then show how social network data can be denoted by a set of sociomatrices.

Sociometry has grown and expanded over the past half century, so that such studies are now usually called simply sociological or occasionally social psychological. The first sociometricians published much of their research in the journal *Sociometry*, which was renamed first *Social Psychology* and then *Social Psychology Quarterly* in the late 1970's. Moreno was *Sociometry*'s founding editor (in 1937). Moreno and other researchers developed a very useful notation for social networks, which we will refer to as sociometric notation. We describe this classic notation in this section.

Sociograms and sociomatrices were first used by Moreno (1934), who demonstrated how they could represent the relational interactions pictured in a sociogram. The focus of Moreno's research, and much of the sociometric literature of the 1930's and 1940's, was how advantageous it was to picture interpersonal interactions using sociograms, even for sets with many actors. In fact, Moreno (see "Emotions Mapped," 1933) aspired to draw a sociometric "map" of New York City, but the best he could do was a sociogram for a community of size 435 (included as a foldout in Moreno 1934). Both Moreno (1934) and Northway (1940) proposed rules for drawing sociograms. These pioneering sociometricians looked for techniques to show the acceptability of each actor relative to the set of actors as a whole and to determine which "choices" were the most important to the group structure. Lindzey and Byrne (1968), building on Moreno's original guidelines, provide a very good discussion of the measurement of relations. Recently, because of innovations in computing, there has been renewed interest in the graphical representations of social network data (Klovdahl 1986).

Moreno actually preferred the use of sociograms to sociomatrices, and had several arguments in print with proponents of sociomatrices, such as Katz. Moreno used his position as editor of *Sociometry* frequently to interject editor's notes into articles in his journal.

Even with the growing interest in figures such as sociograms, researchers were unhappy that different investigators using the same data could produce as many different sociograms (in appearance) as there were investigators. As we have mentioned, the placement of actors and lines in the two-dimensional space is completely arbitrary. Consequently, the use of the sociomatrix to denote social network data increased in the 1940's. The literature in the 1940's presented a variety of methods for analyzing and manipulating sociomatrices (see Dodd 1940; Katz 1947; Festinger 1949; Luce and Perry 1949; and Harary, Norman, and Cartwright 1965). For example, Dodd (1940) described simple algebraic operations for square sociomatrices indexed by the set of actors. He also showed how rows and columns of such matrices could be aggregated to highlight the relationships among sets of actors, rather than the individual actors themselves. Forsyth and Katz (1946) advocated the use of sociomatrices over sociograms to standardize the quantification of social interactions and to represent network data "more objectively" (page 341). This research appears to be the first to focus on derived subgroupings of actors. Katz (1947) proposed a "canonical" decomposition of a sociomatrix to facilitate the comparison of an observed sociomatrix to a target sociomatrix, an idea first proposed by Northway (1940, 1951, 1952). He also showed how sociomatrices could be rearranged using permutation matrices to identify subgroups of actors, and how choices made by a particular actor could be viewed as a multidimensional vector. Festinger (1949) applied matrix multiplication to sociomatrices and described how products of a sociomatrix (particularly squares and cubes) can be used to find cliques or subgroups of similar actors (see also Chabot 1950). Since such powers have simple graph theoretic interpretation (see Chapter 4's discussion of 2- and 3-step walks), this research helped begin the era of graph theoretic approaches to social network analysis. Luce and Perry (1949) and Luce (1950) proposed one of the first techniques to find cliques or subgroups of actors using (for that time) rather sophisticated sociomatrix calculations backed up with an elaborate set of theorems describing the properties and uniqueness of their approach (which was termed *n-clique analysis*; see Chapter 7). Bavelas (1948, 1950) and Leavitt (1951) introduced the notion of centrality (see Chapter 5) into social network analysis. By the end of the decade, researchers had begun to think about electronic calculations for sociometric data (Beum and Criswell 1947; Katz 1950; Beum and Brundage 1950) consisting of a collection of sociomatrices. Research of Katz (1953), MacRae (1960), Wright and Evitts (1961), Coleman (1964), Hubbell (1965), and methods discussed by Mitchell (1969) rely extensively on computers to find various graph theoretic measures. The 1950's and early 1960's became the era of graph theory in sociometry.

The line between sociometric and graph theoretic approaches to social network analysis began to become blurred during the early history of the discipline, as computers began to play a bigger role in data analysis. Sociograms waned in importance as sociomatrices became more popular and as more mathematical and statistical indices were invented that used sociomatrices, much to the dismay of Moreno (1946).

History is certainly on the side of this notational scheme. In fact, most research papers and books on social network methodology begin with the definition of a sociomatrix. Readers who are interested in the topics in Parts II and III will find this notation most useful. For most social network methods, sociometric notation is probably the only notation necessary. It is also the scheme preferred by most network analysis computer programs. It is important to note, however, that sociometric notation can not easily quantify or denote actor attributes, and thus is limited. It is useful when actor attributes are not measured. The relationship between sociometric notation and the more general graph theoretic notation contributes to the popularity of this approach.

As is done throughout this chapter, we split our discussion of sociometric notation and sociomatrices into several parts. We first describe how to construct these two-dimensional sociometric arrays when only one set of actors and one relation is present, and then, when one set of actors and two (or more) relations are measured. Our discussion of two (or more) sets of actors can be found at the end of the chapter.

3.2.1 Single Relation

Let us suppose that we have a single relation measured on one set of g actors in $\mathcal{N} = \{n_1, n_2, \ldots, n_g\}$. We let $\mathscr{X}$ refer to this single valued, directional relation. This relation is measured on the ordered pairs of actors that can be formed from the actors in $\mathcal{N}$.

Consider now the measurements taken on each ordered pair of actors. Define x_{ij} as the value of the tie from the ith actor to the jth actor on the single relation. We now place these measurements into a *sociomatrix*. Rows and columns of this sociomatrix index the individual actors, arranged in identical order. Since there are g actors, the matrix is of size $g \times g$. Sociometric notation uses such matrices to denote measurements on ties.

For the relation $\mathscr{X}$, we define $\mathbf{X}$ as the associated sociomatrix. This sociomatrix has g rows and g columns. The value of the tie from n_i to n_j is placed into the (i, j)th element of $\mathbf{X}$. The entries are defined as:

$$x_{ij} = \text{the value of the tie from } n_i \text{ to } n_j \text{ on relation } \mathscr{X}, \tag{3.1}$$

where i and j $(i \neq j)$ range over all integers from 1 to g. An example will be given shortly. One can think of the elements of $\mathbf{X}$ as the coded values of the relation $\mathscr{X}$. If the relation is dichotomous, then the values for the tie are simply 0 and 1.

Pairs listing the same actor twice, (n_i, n_i), $i = 1, 2, \ldots, g$, are called "self-choices" for a specific relation and are usually undefined. These self-choices lie along the main diagonal of the sociomatrix; consequently, the main diagonal of a sociomatrix is usually full of undefined entries. However, there are situations in which self-choices do make sense. In such cases, the entries $\{x_{ii}\}$ of the sociomatrix are defined. Usually, we will assume undefined sociomatrix diagonals since most methods ignore these elements.

Assume now that this relation is valued and discrete. We will then assume that the possible values for the relation come from the set $\{0, 1, 2, \ldots, C-1\}$, for $C = 2, 3, \ldots$. If the relation is dichotomous, then $C = 2$ possible values. Thus, C is defined as the number of different values the tie can take on. If the relation is valued and discrete, but takes on other than integer values from 0 to $C-1$, then we can easily transform the actual values into the values for this set. For example, if the relation can take on the values $-1, 0, 1$, then we can map -1 to 0, 0 to 1, and $+1$ to 2 (so that $C = 3$). One nice feature of sociometric notation is its ability to handle valued relations.

Since the case of a single relation is just a special case of the multirelational situation, we now turn to this more general case.

3.2.2 Multiple Relations

Suppose that we have R relations $\mathscr{X}_1$, $\mathscr{X}_2$, ..., $\mathscr{X}_R$ measured on a single set of actors. We assume that we have R relations indexed by $r = 1, 2, \ldots, R$. As with a single relation, these relations are valued, and the values for relation $\mathscr{X}_r$ come from the set $\{0, 1, 2, \ldots, C_r - 1\}$.

Consider now the measurements on each possible ordered pair of actors. We define x_{ijr} as the strength of the tie from the ith actor to the jth actor on the rth relation. We now place these measurements into a collection of sociomatrices, one for each relation. Rows and columns of each sociomatrix index the individual actors, arranged in identical order. Thus, the rows and columns of all the sociomatrices are labeled identically. Each matrix is of size $g \times g$.

Consider one of the relations, say $\mathscr{X}_r$, and define $\mathbf{X}_r$ as the sociomatrix associated with this relation. The value of the tie from n_i to n_j is placed into the (i, j)th element of $\mathbf{X}_r$. The entries are defined as:

$$x_{ijr} = \text{the value of the tie from } n_i \text{ to } n_j \text{ on relation } \mathscr{X}_r, \tag{3.2}$$

where i and j $(i \neq j)$ range over all integers from 1 to g, and $r = 1, 2, \ldots, R$. As mentioned, x_{ijr} takes on integer values from 0 to $C_r - 1$. One can think of the elements of $\mathbf{X}_r$ as the coded values of the relation $\mathscr{X}_r$. There are R, $g \times g$ sociomatrices, one for each relation defined for the actors in $\mathscr{N}$. In fact, one can view these R sociomatrices as the layers in a three-dimensional matrix of size $g \times g \times R$. The rows of these sociomatrices index the sending actors, the columns index the receiving actors, and the layers index the relations. Sometimes, this matrix is referred to as a *super-sociomatrix*, representing the information in a multirelational network.

Consider again our example, consisting of a collection of $g = 6$ children and $R = 3$ relations: 1) Friendship at beginning of the school year; 2) Friendship at end of the school year; and 3) Lives near. All three relations are dichotomous, so that $C_1 = C_2 = C_3 = 2$. These three relations are pictured in a single multivariate or multirelational sociogram in Figure 3.2. In Table 3.1 below, we give the three 6×6 dichotomous sociomatrices for the three relations. Note how in Figure 3.2, a "1" in entry (i, j) for the rth sociomatrix indicates that $n_i \rightarrow n_j$ on relation $\mathscr{X}_r$ (or, $n_i \xrightarrow{\mathscr{X}_r} n_j$, for short).

To illustrate, look at the first relation and the first arc in $\mathscr{L}_1$. In Section 3.1, we said that this arc is l_1 = <Allison, Drew>. Allison → Drew is

Table 3.1. *Sociomatrices for the six actors and three relations of Figure 3.2*

Friendship at Beginning of Year

	Allison	Drew	Eliot	Keith	Ross	Sarah
Allison	-	1	0	0	1	0
Drew	0	-	1	0	0	1
Eliot	0	1	-	0	0	0
Keith	0	0	0	-	1	0
Ross	0	0	0	0	-	1
Sarah	0	1	0	0	0	-

Friendship at End of Year

	Allison	Drew	Eliot	Keith	Ross	Sarah
Allison	-	1	0	0	1	0
Drew	0	-	1	0	1	1
Eliot	0	0	-	0	1	0
Keith	0	1	0	-	1	0
Ross	0	0	0	1	-	1
Sarah	0	1	0	0	0	-

Lives Near

	Allison	Drew	Eliot	Keith	Ross	Sarah
Allison	-	0	0	0	1	1
Drew	0	-	1	0	0	0
Eliot	0	1	-	0	0	0
Keith	0	0	0	-	1	1
Ross	1	0	0	1	-	1
Sarah	1	0	0	1	1	-

represented by the arc l_1. Thus, there is an arc from Allison to Drew in the sociogram for the first relation, indicating that Allison chooses Drew as a friend at the beginning of the school year. The first entry in $\mathscr{L}_1$ is exactly this arc. This arc is how this tie is denoted by graph theoretic notation. Consider now how this single tie is coded with sociometric notation. Consider the first sociomatrix in Table 3.1. Consider the entry which quantifies Allison (n_1) as a sender (the first row) and Drew (n_2) as a receiver (the second column) on relation $\mathscr{X}_1$. This entry is in the $(1, 2)$ cell of this sociomatrix, and contains a 1 indicating that

$$\begin{aligned} x_{121} &= \text{the value of the tie from } n_1 \text{ to } n_2 \text{ on relation } \mathscr{X}_1 \\ &= 1. \end{aligned}$$

Note also that $x_{211} = 0$, indicating that Drew does not choose Allison as a friend at the beginning of the school year; that is, Drew $\not\rightarrow$ Allison. This friendship is clearly one-sided, and is not reciprocated.

As one can see, sociometric notation is simple, once one gets used to reading information from two-dimensional sociomatrices. Also note how the diagonals of all three sociomatrices in Table 3.1 are undefined — by design, children are not allowed to choose themselves as friends, and we do not record whether a child lives near himself or herself.

These sociomatrices are the adjacency matrices for the two directed graphs and one undirected graph for the three dichotomous relations. The graphs and the sociomatrices represent exactly the same information. In graph theoretic notation, there are two sets of arcs and one set of lines, $\mathscr{L}_1$, $\mathscr{L}_2$, $\mathscr{L}_3$, which list the ordered pairs of children that are tied for the first two relations and the pairs of children that are tied for the third. If an ordered pair is included in the first or second $\mathscr{L}$ set, then there is an arc drawn from the first child in the pair (the sender) to the second (the receiver). And if an unordered pair of actors is included in the third line set, then there is a line between the two children in the pair. In sociometric notation, the entry in the corresponding cell of the sociomatrix is unity.

We also want to note that the third relation in this network data set is nondirectional; that is, there is a line from n_i to n_j whenever there is a line from n_j to n_i, and vice versa. Note how we were able to code this relation in the sociomatrix given in Table 3.1. Also note that the sociomatrix for a nondirectional relation is symmetric; that is, $x_{ij} = x_{ji}$. One very nice feature of sociometric notation is that it can easily handle both directional and nondirectional relations.

3.2.3 Summary

As we have stated in this section, sociometric notation is the oldest, and perhaps the easiest, way to denote the ties among a set of actors. A single two-dimensional sociomatrix is defined for each relation, and the entries of this matrix code the ties between pairs of actors. Generalizing to valued relations is also easy — the entries in a sociomatrix are the values of the ties, not simply 0's and 1's.

Sociometric notation is very common, the notation of choice for network computing, and will be our first choice of a notational scheme throughout this book. However, as we have mentioned, there are network data sets for which sociometric notation is more difficult to use — specifically, those which contain information on the attributes of the actors. For example consider our second-graders. If we knew their ethnicity (coded on some nominal scale), it would be difficult to include this information in the three sociomatrices (but see Frank and Harary 1982, for an alternative representational scheme).

To conclude, we will frequently use sociomatrices to present network data. These arrays are very convenient (and space-saving!) devices to denote network data sets.

3.3 ○Algebraic Notation

Let us now focus on relations in multirelational networks. In order to present algebraic methods and models for multiple relations (such as relational algebras) in Chapters 11 and 12, it is useful to employ a notation that is different from, though consistent with, the sociometric and graph theoretic notations that we have just discussed. We will refer to this scheme as *algebraic notation.* Algebraic notation is most useful for multirelational networks since it easily denotes the "combinations" of relations in these networks. However, it can also be used to describe data for single relational networks.

There are two major differences between algebraic notation and sociometric notation. First, one refers to relations with distinct capital letters, rather than with subscripted $\mathscr{X}$'s. For example, we could use F to denote the relation "is a friend of" and E for the relation "is an enemy of." Second, we will record the presence of a tie from actor i to actor j on relation F as iFj. This is a shorthand for the sociometric and graph theoretic notation. Rather than indicating ties as $i \rightarrow j$, we will replace the $\rightarrow$ with the letter label for the relation.

In general, $x_{ijF} = 1$ if $n_i \rightarrow n_j$ on the relation labeled $\mathscr{X}_F$ (or F for short). This tie will be denoted by $i \xrightarrow{F} j$, or shortened even further to iFj. This latter notation, iFj, is algebraic.

Referring to our example, we label the relation "is a friend of at the beginning of the school year" as F. We would record the tie implied by "child i chooses child j as a friend at the beginning of the school year" as iFj. In sociometric notation, iFj means that $x_{ijF} = 1$, and implies that there is a "1" in the cell at row i and column j of the sociomatrix for this relation.

Algebraic notation is especially useful for dichotomous relations, since it codes the presence of ties on a given relation. Extensions to valued relations can be difficult. However, the limitation to dichotomous relations presents no problem for us, since the models that use algebraic notation are specific to dichotomous relations. The advantages of this notation are that it allows us to distinguish several distinct relations using letter designations, and to record *combinations* of relations, such as "friends' enemy," or "mother's brother," or a "friend's neighbor." Unfortunately, this notational scheme can not handle valued relations or actor attributes.

3.4 ○Two Sets of Actors

A network may include two sets of actors. Such a network is a two-mode network, with each set of actors constituting one of the modes. A researcher studying such a network might focus on how the actors in one set relate to each other, how the actors in the other set relate to each other, and/or how actors in one set relate to the actors in the other set. In this situation, we need to distinguish between the two sets of actors and the different types of ties. We note that relations defined on two sets of actors often yield complicated network data sets. It is thus quite complicated to give "hard-and-fast" notation rules to apply to every and all situations. We recommend that for multirelational data sets one make an inventory of measured relations and modify the rules given below to apply to the situation at hand.

There are many social networks that involve two sets of actors. For example, we might have a collection of teachers and students who are interacting with each other. Consider the relations "is a student of" and "attends faculty meetings together." The relation "is a student of" can only exist between a student and a teacher. The relation "attends faculty meetings together" is defined only for pairs of teachers.

We will call the first actor in the pair the *sender* and the second actor the *receiver*. Other authors have called these actors *originators* and *recipient*, or simply, *actors* and *partners*. With this understanding, we can distinguish between the two actors in the pair. If the relation is defined on a single set of actors, both actors in the pair can be senders and both can be receivers. The interesting "wrinkle" that arises if there are two sets of actors is that the senders might come only from the first set and the receivers only from the second.

We will let $\mathscr{N}$ refer to the first set of actors and $\mathscr{M}$ refer to the second set. The set $\mathscr{N}$ contains g actors and the second set $\mathscr{M}$ contains h actors. The set $\mathscr{M}$ contains elements $\{m_1, m_2, \ldots, m_h\}$, so that m_i is a typical actor in the second set. Further, there are $\binom{h}{2}$ dyads that can be formed from actors in $\mathscr{M}$.

In this section, we will first discuss the two types of pairs that can arise when relations are measured on two (or even more) sets of actors. We present only sociometric notation, since it is sufficient.

To illustrate the notation, we return to our collection of six second-grade children, and now consider a second set of actors, $\mathscr{M}$, consisting of $h = 4$ adults. We define m_1 = Mr. Jones, m_2 = Ms. Smith, m_3 = Mr. White, and m_4 = Ms. Davis. In total, we have ten actors, which are grouped into these two sets. Considering just the actors in $\mathscr{M}$, there are $4(4-1)/2 = 6$ additional unordered pairs.

3.4.1 ⊗Different Types of Pairs

With two sets of actors, there can be two types of pairs — those that consist of actors from the same set and those that consist of actors from different sets. We will call the former *homogeneous* and the latter *heterogeneous*. Thus, in homogeneous pairs the senders and receivers are from the same set, while in heterogeneous pairs actors are from different sets. We discuss each of these types, beginning with homogeneous pairs.

We can further distinguish between two kinds of homogeneous pairs by noting that there are two sets from which the actors can come. The two kinds of homogeneous pairs are:

- Sender and Receiver both belong to $\mathscr{N}$
- Sender and Receiver both belong to $\mathscr{M}$

In a data set with just one set of actors, the pairs are all homogeneous. However, when there are two sets of actors, there are two kinds of homogeneous pairs.

Of more interest when there are two sets of actors are the pairs that contain one actor from each set. These heterogeneous pairs are also of two kinds, depending on the sets to which the sender and receiver belong. Assuming the relation for the heterogeneous pairs is directional, the originating actor must belong to a different set than the receiving actor. Since there are two sets of actors, we get two kinds of heterogeneous pairs:

- Sender belongs to $\mathscr{N}$ and Receiver belongs to $\mathscr{M}$
- Sender belongs to $\mathscr{M}$ and Receiver belongs to $\mathscr{N}$

It is important to distinguish between these two collections of heterogeneous pairs. Relations defined on the first collection of pairs can be quite different from those defined on the second. For example, if

$\mathcal{N}$ is a set of major corporations in a large city and $\mathcal{M}$ is a set of non-profit organizations (such as churches, arts organizations, charitable institutions, etc.), then we could study how the corporations in $\mathcal{N}$ make charitable contributions to the non-profits in $\mathcal{M}$. Such a relation would not be defined for the other collection of heterogeneous pairs, since it is virtually impossible for non-profits to contribute money to the welfare of the corporations.

3.4.2 ○*Sociometric Notation*

We now turn our attention to sociometric notation and sociomatrices for the relations defined for both homogeneous *and* heterogeneous pairs. The notation will have to allow for the fact that the sending and receiving actors could come from different sets. We assume that we have a number of relations. The measurements for a specific relation can be placed into a sociomatrix, and there is one sociomatrix for each relation.

A sociomatrix is indexed by the set of originating actors (for its rows) and the set of receiving actors (for its columns) and gives the values of the ties from the row actors to the column actors. If the relation is defined for actors from different sets, then in general, its sociomatrix will not be square. Rather, it will be *rectangular.*

Let us pick one of the relations, say $\mathscr{X}_r$, and suppose that it is defined on a collection of heterogeneous pairs in which the originating actor is from $\mathcal{N}$ and the receiving actor is from $\mathcal{M}$. The sociomatrix $\mathbf{X}_r$, giving the measurements on $\mathscr{X}_r$, has dimensions $g \times h$. The (i, j)th cell of this matrix gives the measurement on this rth relation for the pair of actors (n_i, m_j). The (i, j)th entry of the sociomatrix $\mathbf{X}_r$ is defined as:

$$\begin{aligned} x_{ijr} \quad = \quad & \text{the value of the tie from } n_i \text{ to } m_j \\ & \text{on the relation } \mathscr{X}_r. \end{aligned} \tag{3.3}$$

The actor index i ranges over all integers from 1 to g, while j ranges over all integers from 1 to h, and $r = 1, 2, \ldots, R$. As with relations defined on a single set of actors, x_{ijr} takes on integer values from 0 to $C_r - 1$.

Here, i can certainly equal j, since these two indices refer to different sets. The value of x_{iir} is meaningful.

When there are two sets of actors, there are four possible types of sociomatrices, each of which might be of a different size. The rows and columns of the sociomatrices will be labeled by the actors in the sets involved: the rows for the sending actor set and the columns for

Table 3.2. *The sociomatrix for the relation "is a student of" defined for heterogeneous pairs from $\mathcal{N}$ and $\mathcal{M}$*

		$\mathcal{M}$			
		Mr. Jones	Ms. Smith	Ms. Davis	Mr. White
	Allison	1	0	0	0
	Drew	0	1	0	0
$\mathcal{N}$	Eliot	0	0	1	0
	Keith	0	0	0	1
	Ross	0	0	1	0
	Sarah	0	1	0	0

the receiving actor set. We will denote the sociomatrices by using their sending and receiving actor sets, so, for example, the sociomatrix $\mathbf{X}^{\mathcal{N}\mathcal{M}}$ contains measurements on a relation defined from actors in $\mathcal{N}$ to actors in $\mathcal{M}$. These sociomatrices and their sizes are:

- $\mathbf{X}_r^{\mathcal{N}}$, dimensions $= g \times g$
- $\mathbf{X}_r^{\mathcal{M}}$, dimensions $= h \times h$
- $\mathbf{X}_r^{\mathcal{N}\mathcal{M}}$, dimensions $= g \times h$
- $\mathbf{X}_r^{\mathcal{M}\mathcal{N}}$, dimensions $= h \times g$

The second two types are, in general, rectangular. As always, in each sociomatrix, x_{ijr} is the value of the tie from actor i to actor j on the rth relation of that particular type.

Clearly, this notational scheme can accommodate multiple relations. However, since there may be a different number of relations defined for the four different types of pairs of actors, there may be different numbers of sociomatrices of each type.

To illustrate, consider an example with two sets of actors: students and teachers. Suppose there are four adults, second-grade teachers at the elementary school that is attended by six children. Define a relation, "is a student of." This relation is defined for heterogeneous pairs of actors for which the sender belongs to $\mathcal{N}$ and the receiver belongs to $\mathcal{M}$; that is, a child "is a student of" an adult teacher, but not vice versa. Table 3.2 gives the sociomatrix for the two-mode relation "is a student of" from our network of second-grade children. This relation is defined for the heterogeneous pairs consisting of a child as the sender and an adult as a receiver. This is a dichotomous relation ($C = 2$), and is measured on the $6 \times 4 = 24$ heterogeneous pairs of children and teachers.

Note that there is only one 1 in every row of this matrix, since a child can have only one teacher. The entries in a specific column give the children that are taught by each teacher. Note how easily this array codes the information in the directional relation between two sets of actors. It is important to note that with sociometric notation all we need is one sociomatrix (with the proper dimensions) for each relation.

3.5 Putting It All Together

We conclude this chapter by pulling together all three notations into a single, more general framework. To begin, we note that the collection of actors, the relational information on pairs of actors, and possible attributes of the actors constitute a collection of data that can be referred to as a *social relational system.* Such a system is a conceptualization of the actors, pairs, relations, and attributes found in a social network.

As we have shown in this chapter, the data for a social relational system can be denoted in a variety of ways. It is important to stress that when dichotomous relations are considered, the three notational systems discussed in this chapter are capable of representing the entire data set.

We will use the symbols "$n_i \to n_j$" as shorthand notation for n_i "chooses" n_j on the single relation in question; that is, the arc from n_i to n_j is contained in the set $\mathscr{L}$, so that there is a tie present for the ordered pair $< n_i, n_j >$. If this arc is an element of $\mathscr{L}$, then there is a directed line from node i to node j in the directed graph or sociogram representing the relationships between pairs of actors on the relation. Sometimes we will replace "$n_i \to n_j$" with "$i \to j$" if no confusion could arise. With algebraic notation, if we label this relation by, say, F, we can also state that iFj. And with sociometric notation, we record this tie as $x_{ij} = 1$ in the proper sociomatrix.

As we have mentioned in our discussion of graph theoretic notation, if one has a single set of g actors, $\mathscr{N}$, then there are $g(g-1)$ ordered pairs of actors. In addition to $\mathscr{N}$, the set $\mathscr{L}$ contains the collection of ordered pairs of actors for which ties are present.

Some social network methodologists refer to the set of actors and the set of arcs as the *algebraic structure* $S = <\mathscr{N}, \mathscr{L}>$ (Freeman 1989). S is the standard representation of the simplest possible social network. For us, this is the graph theoretic representation.

One can define a graph from S by stating that the directed graph $\mathscr{G}_d$ is the ordered pair $<\mathscr{N}, \mathscr{L}>$, where the elements of $\mathscr{N}$ are nodes in the graph, and the elements of $\mathscr{L}$ are the ordered pairs of nodes for which there is a tie from n_i to n_j ($n_i \to n_j$).

Nodes and arcs are the basic building blocks for graph theoretic notation. To relate these concepts to the elements of sociometric notation, we consider again the collection of all ordered pairs of actors in $\mathcal{N}$. Sometimes this collection is denoted $\mathcal{N} \times \mathcal{N}$, a Cartesian product of sets. We define a binary quantity x_{ij} to be equal to 1 if the ordered pair $< n_i, n_j >$ is an element of $\mathcal{L}$ (that is, if there is a tie from n_i to n_j) and equal to 0 if the ordered pair is not an element of $\mathcal{L}$. This quantity is a mapping from the elements of the collection of ordered pairs to the set containing just 0 and 1. These quantities are exactly the elements of the binary $g \times g$ sociomatrix $\mathbf{X}$.

A relation is the collection of all ordered pairs for which $n_i \rightarrow n_j$. It is thus a subset of $\mathcal{N} \times \mathcal{N}$. In algebraic notation, capital letters (such as F) are used to refer to specific relations and to denote which ties are present. A relation is thus the set of all pairs of actors for which $n_i \rightarrow n_j$, or $x_{ij} = 1$, or iFj.

Thus, one can see the equivalence between the graph theoretic notation, and the sociometric notation (built on sociomatrices), and the algebraic notation (dependent on relations such as F). Freeman (1989) views the triple consisting of the algebraic structure S, the directed graph or sociogram $\mathcal{G}_d$, and the adjacency matrix or sociomatrix $\mathbf{X}$ as a social network:

$$\mathcal{S} = < S, \ \mathcal{G}_d, \ \mathbf{X} > .$$

This triple provides a nice abstract definition of the central concept of this book. And, it shows how these notational schemes are usually viewed together as providing the three essential components of the simplest form of a social network:

- A set of nodes and a set of arcs (from graph theoretic notation)
- A sociogram or graph (produced from the sets of nodes and arcs)
- A sociomatrix (from sociometric notation)

It is important to note that most of the generalizations of this simple social network $\mathcal{S}$, such as to valued relations, multiple relations, more than one set of actors, and relations measured over time, can be viewed in just the same way as the situation described here (single dichotomous relation measured on a single set of actors). The only wrinkle is that actor attributes are not easily quantified by using these concepts. The best one can do is to define a new matrix, $\mathbf{A}$, of dimensions (number of actors) $\times$ (number of attributes) to hold the measurements on the attribute variables. One could even include this information in the social network definition, so that a more complicated social network is $\mathscr{S} = < S, \mathscr{G}_d, \mathbf{X}, \mathbf{A} >$.

Lastly, we should note that nowhere in this chapter did we discuss affiliation relations. We have introduced affiliation networks in Chapter 2, and will defer a mathematical description until Chapter 8.

4

Graphs and Matrices

by Dawn Iacobucci

This chapter presents the terminology and concepts of *graph theory*, and describes basic matrix operations that are used in social network analysis. Both graph theory and matrix operations have served as the foundations of many concepts in the analysis of social networks (Hage and Harary 1983; Harary, Norman, and Cartwright 1965). In this chapter, the notation presented in Chapter 3 is used, and more concepts and ideas from graph theory are described and illustrated with examples. The topics covered in this chapter are important for the methods discussed in the remaining chapters of the book, but they are especially important in Chapter 5 (Centrality, Prestige, and Related Actor and Group Measures), Chapter 6 (Structural Balance, Clusterability, and Transitivity), Chapter 7 (Cohesive Subgroups), and Chapter 8 (Affiliations, Co-memberships, and Overlapping Subgroups).

We start this chapter with a discussion of some reasons why graph theory and graph theoretic concepts are important for social network analysis. We then define a graph for representing a nondirectional relation. We begin with simple concepts, and progressively build on these to achieve more complicated, and more interesting, graph theoretic concepts. We then define and discuss directed graphs, for representing directional relations. Again, we begin with simple directed graph concepts and build to more complicated ideas. Following this, we discuss signed and valued graphs. We then define and discuss hypergraphs, which are used to represent affiliation networks. In the final section of this chapter we define and illustrate basic matrix operations that are used in social network analysis, and show how many of these matrix operations can be used to study the graph theoretic concepts discussed in the earlier sections of this chapter.

4.1 Why Graphs?

Graph theory has been useful in social network analysis for many reasons. Among these reasons are the following (see Harary, Norman, and Cartwright 1965, page 3). First, graph theory provides a vocabulary which can be used to label and denote many social structural properties. This vocabulary also gives us a set of primitive concepts that allows us to refer quite precisely to these properties. Second, graph theory gives us mathematical operations and ideas with which many of these properties can be quantified and measured (see Freeman 1984; Seidman and Foster 1978b). Last, given this vocabulary and these mathematics, graph theory gives us the ability to prove theorems about graphs, and hence, about representations of social structure. Like other branches of mathematics, graph theory allows researchers to prove theorems and deduce testable statements. However, as Barnes and Harary (1983) have noted, "Network analysts ... make too little use of the *theory* of graphs" (page 235). Although the representation of a graph and the vocabulary of graph theory are widely used by social network researchers, the theorems and derivations of graph theory are less widely used by network methodologists. Some notable exceptions include the work of Davis, Everett, Frank, Hage, Harary, Johnsen, Peay, Roberts, and Seidman, among others.

In addition to its utility as a mathematical system, graph theory gives us a representation of a social network as a model of a social system consisting of a set of actors and the ties between them. By *model* we mean a simplified representation of a situation that contains some, but not all, of the elements of the situation it represents (Roberts 1976; Hage and Harary 1983). When a graph is used as a model of a social network, points (called *nodes*) are used to represent actors, and lines connecting the points are used to represent the ties between the actors. In this sense, a graph is a model of a social network, in the same way that a model train set is a model of a railway system.

Graphs have been widely used in social network analysis as a means of formally representing social relations and quantifying important social structural properties, beginning with Moreno (1934), and developed further by Harary (Harary 1959a; Harary 1959b; Harary 1969; Hage and Harary 1983; Harary, Norman, and Cartwright 1965) and others (for example, Frank 1971; Seidman and Foster 1978a, 1987b; Foster and Seidman 1982, 1983, 1984). Graph theory has been used heavily in anthropology (Mitchell 1980; Hage 1973, 1976a, 1976b, 1979; Hage and Harary 1983; Abell 1970; Barnes 1969b; Barnes and Harary 1983; Zachary 1977), social psychology (Heider 1944, 1946, 1958; Davis 1967; Bavelas 1948, 1950; Leavitt 1951; Freeman 1977, 1979; Freeman, Roeder, and Mulholland 1980), communications, business, organizational research, and geography (Pitts 1965, 1979).

The visual representation of data that a graph or sociogram offers often allows researchers to uncover patterns that might otherwise go undetected (Moreno 1934; Hoaglin, Mosteller, and Tukey 1985; Tukey 1977; Velleman and Hoaglin 1981).

Matrices are an alternative way to represent and summarize network data. A matrix contains exactly the same information as a graph, but is more useful for computation and computer analysis. Matrix operations are widely used for definition and calculation in social network analysis, and are the primary representation for most computer analysis packages (*GRADAP, UCINET, STRUCTURE, SNAPS, NEGOPY*). However, only the program *GRADAP* is explicitly graph theoretic.

We will illustrate the graph theoretic concepts discussed in this chapter on small, simple social networks. Most of these examples will consist of hypothetical data created to demonstrate specific properties of graphs. We will also refer to the data collected by Padgett on the marital alliances between sixteen families in 15th century Florence, Italy.

In the following section, we describe properties of *graphs*, where a line between two nodes is *nondirectional*. Graphs are used for representing nondirectional relations. Following the discussion of graphs, we describe properties of *directed graphs*, where a line is directed from one node to another. Directed graphs, or *digraphs*, are used for representing directional relations, where the tie has an origin and a destination.

4.2 Graphs

A graph is a model for a social network with an undirected dichotomous relation; that is, a tie is either present or absent between each pair of actors. Nondirectional relations include such things as co-membership in formal organizations or informal groups, some kinship relations such as "is married to," "is a blood relative of," proximity relations such as "lives near," and interactions such as "works with." In a graph, *nodes* represent actors and *lines* represent ties between actors. In graph theory, the nodes are also referred to as *vertices* or *points*, and the lines are also known as *edges* or *arcs*.

A *graph* $\mathscr{G}$ consists of two sets of information: a set of *nodes*, $\mathscr{N} = \{n_1, n_2, \ldots, n_g\}$, and a set of lines, $\mathscr{L} = \{l_1, l_2, \ldots, l_L\}$ between pairs of nodes. There are g nodes and L lines. In a graph each line is an unordered pair of distinct nodes, $l_k = (n_i, n_j)$. Since lines are unordered pairs of nodes, the line between nodes n_i and n_j is identical to the line between nodes n_j and n_i ($l_k = (n_i, n_j) = (n_j, n_i)$). We will exclude the possible line between a node and itself, (n_i, n_i). Such lines are called *loops* or *reflexive ties*. Also, we do not allow an unordered pair of nodes to be included more than once in the set of lines. Thus, there can be no more than one line between a pair of nodes. A graph that has no loops and includes no more than one line between a pair of nodes is called a *simple graph*. Unless we note otherwise, the graphs that we consider in this chapter are simple graphs.

In a graph of a social network with a single nondirectional dichotomous relation, the nodes represent actors, and the lines represent the ties that exist between pairs of actors on the relation. A line $l_k = (n_i, n_j)$ is included in the set of lines, $\mathscr{L}$, if there is a tie present between the two actors in the network who are represented by nodes n_i and n_j in the graph.

Taken together, the two sets of information (nodes and lines) may be used to refer formally to a graph in terms of its node set and its line set. Thus we can denote a graph with node set $\mathscr{N}$ and line set $\mathscr{L}$ as $\mathscr{G}(\mathscr{N}, \mathscr{L})$. However, when there is no ambiguity about the node set and the line set, we will refer to a graph simply as $\mathscr{G}$.

Two nodes, n_i and n_j, are *adjacent* if the line $l_k = (n_i, n_j)$ is in the set of lines $\mathscr{L}$. A node is *incident* with a line, and the line is incident with the node, if the node is one of the unordered pair of nodes defining the line. For example, nodes n_1 and n_2 are incident with line $l_1 = (n_1, n_2)$. Each line is incident with the two nodes in the unordered pair that define the line.

A graph that contains only one node is *trivial*; all other graphs are nontrivial. A graph that contains g nodes and no lines ($L = 0$) is *empty*. Trivial and empty graphs are of little substantive interest. In social networks, these graphs would correspond to a network consisting of only one actor (the trivial graph) and a network consisting of more than one actor, but no ties between the actors (the empty graph).

A graph $\mathscr{G}(\mathscr{N}, \mathscr{L})$ can also be presented as a diagram in which points depict nodes, and a line is drawn between two points if there is a line between the corresponding two nodes in the set of lines, $\mathscr{L}$. The location

	Actor	*Lives near:*
n_1	Allison	Ross, Sarah
n_2	Drew	Eliot
n_3	Eliot	Drew
n_4	Keith	Ross, Sarah
n_5	Ross	Allison, Keith, Sarah
n_6	Sarah	Allison, Keith, Ross

$l_1 = (n_1, n_5)$
$l_2 = (n_1, n_6)$
$l_3 = (n_2, n_3)$
$l_4 = (n_4, n_5)$
$l_5 = (n_4, n_6)$
$l_6 = (n_5, n_6)$

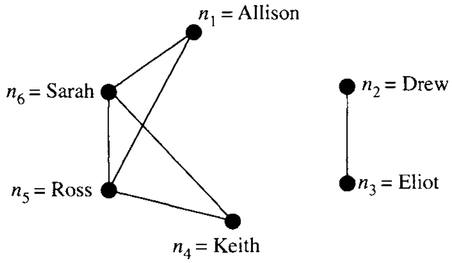


Fig. 4.1. Graph of "lives near" relation for six children

of points on the page is arbitrary, and the length of the lines between points is meaningless. The only information in the graph is the set of nodes and presence or absence of lines between pairs of nodes. In social network analysis, such a diagram is frequently referred to as a *sociogram*.

An example of a graph is given in Figure 4.1. We begin with a small graph so that all of its elements may be easily identified. The sets of nodes and lines are also listed. In this example, we can take the six nodes to represent the six children and the undirected relation "lives near," discussed in Chapter 3. In this example there are $g = 6$ nodes and $L = 6$ lines. A line between two nodes indicates that the children represented by these nodes live near each other. For example, Sarah, n_6, and Allison, n_1, live near each other so the line (n_1, n_6) is included in the set of lines. Allison and Eliot, n_3, do not live near each other, so the line (n_1, n_3) is not in the set of lines.

Social networks can be studied at several levels: the actor, pair or dyad, triple or triad, subgroup, and the group as a whole. In graph theoretic terms, these levels correspond to different *subgraphs*. Many social network methods consider subgraphs contained in a graph. For example, dyads and triads are (very small) subgraphs.

4.2.1 Subgraphs, Dyads, and Triads

Subgraphs. A graph $\mathscr{G}_s$ is a *subgraph* of $\mathscr{G}$ if the set of nodes of $\mathscr{G}_s$ is a subset of the set of nodes of $\mathscr{G}$, and the set of lines in $\mathscr{G}_s$ is a subset of the lines in the graph $\mathscr{G}$. If we denote the nodes in $\mathscr{G}_s$ as $\mathscr{N}_s$ and the lines in $\mathscr{G}_s$ as $\mathscr{L}_s$, then $\mathscr{G}_s$ is a subgraph of $\mathscr{G}$ if $\mathscr{N}_s \subseteq \mathscr{N}$ and $\mathscr{L}_s \subseteq \mathscr{L}$. All lines in $\mathscr{L}_s$ must be between pairs of nodes in $\mathscr{N}_s$. However, since $\mathscr{L}_s$ is a subset of $\mathscr{L}$, there may be lines in the graph between pairs of nodes in the subgraph that are not included in the set of lines in the subgraph.

Figure 4.2 gives an example of a graph and some of its subgraphs. In $\mathscr{G}$, the set of nodes consists of $\mathscr{N} = \{n_1, n_2, n_3, n_4, n_5\}$ and $\mathscr{L} = \{l_1, l_2, l_3, l_4\}$. In the subgraph in Figure 4.2b the set of nodes is $\mathscr{N}_s = \{n_1, n_3, n_4\}$ and the set of lines is $\mathscr{L}_s = \{l_2\}$. Notice that the subgraph does not include the line $l_4 = (n_3, n_4)$.

Any generic subgraph may not include all lines between the nodes in the subgraph. There are (at least) two special kinds of subgraphs that can be derived from a graph. One can take a subset of nodes and consider all lines that are between the nodes in the subset. Such a subgraph is *node-generated*, since the subset of nodes has produced the subgraph. Or, one can take a set of subset of lines, and consider all nodes that are incident with the lines in the subset. Such a subgraph is *line-generated*. We discuss each of these below.

Node- and Line-Generated Subgraphs. First consider node-generated subgraphs. A subgraph, $\mathscr{G}_s$, is *generated by a set of nodes*, $\mathscr{N}_s$, if $\mathscr{G}_s$ has node set $\mathscr{N}_s$, and line set $\mathscr{L}_s$, where the set of lines, $\mathscr{L}_s$, includes *all* lines from $\mathscr{L}$ that are between pairs of nodes in $\mathscr{N}_s$. Whereas a subgraph does not necessarily include all of the lines from $\mathscr{L}$ that are between nodes in $\mathscr{N}_s$, a subgraph generated by node set $\mathscr{N}_s$ must include all lines from $\mathscr{L}$ that are present between pairs of nodes in $\mathscr{N}_s$.

In social network analysis, a node-generated subgraph results if the researcher considers only a subset of the g members of the network. Some

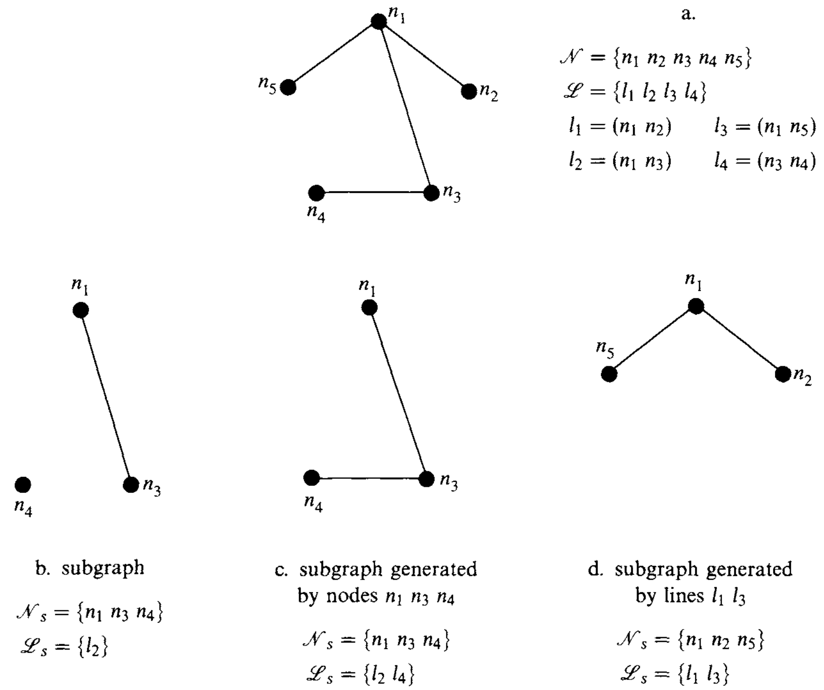


Fig. 4.2. Subgraphs of a graph

relational data might be missing for some of the network members, and thus the researcher can only study ties among the remaining actors. In a longitudinal study in which a network is studied over time, some actor, or subset of actors, might leave the network. Analyses of the network might have to be restricted to the subset of actors for whom data are available for all time points. Node-generated subgraphs are widely used in the analysis of cohesive subgroups in networks (see Chapter 7). These methods focus on subsets of actors among whom the ties are relatively strong, numerous, or close.

Now consider line-generated subgraphs. A subgraph, $\mathscr{G}_s$, is *generated by a set of lines*, $\mathscr{L}_s$, if $\mathscr{G}_s$ has line set $\mathscr{L}_s$, and node set $\mathscr{N}_s$, where the set of nodes, $\mathscr{N}_s$, includes all nodes from $\mathscr{N}$ that are incident with lines in $\mathscr{L}_s$. Figure 4.2c shows the subgraph generated by the set of nodes $\mathscr{N}_s = \{n_1, n_3, n_4\}$. In this subgraph both lines $l_2 = (n_1, n_3)$ and $l_4 = (n_3, n_4)$ are included, since a node-generated subgraph includes all lines between the nodes in the graph. Figure 4.2d shows the subgraph generated by the set of lines $\mathscr{L}_s = \{l_1, l_3\}$. The set of nodes in this subgraph includes all nodes incident with lines l_1 and l_3, so $\mathscr{N}_s = \{n_1, n_2, n_5\}$.

So there are two special kinds of subgraphs that we will consider. Throughout the book, most of the subgraphs we consider will be node-generated subgraphs.

An important feature of a subgraph is whether it is *maximal* with respect to some property. A subgraph is maximal with respect to a given property if that property holds for the subgraph $\mathscr{G}_s$, but does not hold if any node or nodes are added to the subgraph. We will return to this property and illustrate it later in this chapter.

Dyads. A dyad, representing a pair of actors and the possible tie between them, is a (node-generated) subgraph consisting of a pair of nodes and the possible line between the nodes. In a graph an unordered pair of nodes can be in only one of two states: either two nodes are adjacent or they are not adjacent. Thus, there are only two dyadic states for an undirected relation represented as a graph; either the actors in the dyad have a tie present, or they do not.

Triads. Triadic analysis is also based on subgraphs, where the number of nodes is three. A triad is a subgraph consisting of three nodes and the possible lines among them. In a graph, a triad may be in one of four possible states, depending on whether, zero, one, two, or three lines are present among the three nodes in the triad. These four possible triadic states are shown in Figure 4.3.

There has been much theoretical research on triads. For example, Granovetter (1973) refers to the triad with two lines present and one line absent as the *forbidden triad.* He argues that if lines represent strong ties between actors, then if actor i has a strong tie with actor j, and actor j in turn has a strong tie with actor k, it is unlikely that the tie between actor i and actor k will be absent. This type of triad, with only two lines, is forbidden in Granovetter's model.

Both dyads and triads are node-generated subgraphs, since they are defined as a subset of nodes and all lines between pairs of nodes in the subset.

We now consider properties of nodes and graphs that can be defined using the concepts of adjacency and incidence for the nodes and lines in a graph.

Fig. 4.3. Four possible triadic states in a graph

4.2.2 Nodal Degree

The *degree* of a node, denoted by $d(n_i)$, is the number of lines that are incident with it. Equivalently, the degree of a node is the number of nodes adjacent to it. The degree of a node is a count that ranges from a minimum of 0, if no nodes are adjacent to a given node, to a maximum of $g - 1$, if a given node is adjacent to all other nodes in the graph. A node with degree equal to 0 is called an *isolate.*

The degree of a node, $d(n_i)$, may be obtained by counting the number of lines incident with it. In the example in Figure 4.1, the degrees of the nodes are: $d(n_1) = 2, d(n_2) = 1, d(n_3) = 1, d(n_4) = 2, d(n_5) = 3$, and $d(n_6) = 3$.

Degrees are very easy to compute, and yet can be quite informative in many applications. For example, if we observe children playing together, and represent children by nodes, and instances of pairs of children playing by lines in a graph, then a node with a small degree would indicate a child who played with few others, and a node with a large degree would indicate a child who played with many others. Or, in the study of Padgett's Florentine families, and the relation of marriage, a node with a large degree represents a family that has many marital ties to other families in the network. The degree of a node is a measure of the "activity" of the actor it represents, and is the basis for one of the centrality measures that we discuss in Chapter 5.

In many applications, it is informative to summarize the degrees of all the actors in the network. The mean nodal degree is a statistic that reports the average degree of the nodes in the graph. Denoting the mean degree as $\bar{d}$, we have

$$\bar{d} = \frac{\sum_{i=1}^{g} d(n_i)}{g} = \frac{2L}{g}. \tag{4.1}$$

One might also be interested in the variability of the nodal degrees. If all the degrees of all of the nodes are equal, the graph is said to be *d-regular*, where d is the constant value for all the degrees ($d(n_i) = d$, for all i and some value d). d-regularity can be thought of as a measure of uniformity. We will discuss d-regularity in more detail below in the context of directed graphs. If a graph is not d-regular, the nodes differ in degree. The variance of the degrees, which we denote by S_D^2, is calculated as:

$$S_D^2 = \frac{\sum_{i=1}^{g}(d(n_i) - \bar{d})^2}{g}. \tag{4.2}$$

A graph that is d-regular has $S_D^2 = 0$. Variability in nodal degrees means that the actors represented by the nodes differ in "activity," as measured by the number of ties they have to others. The variability of nodal degrees is one measure of graph centralization that we discuss in Chapter 5.

The nodal degrees are an important property of a graph, and we will often want to control for or condition on the set of nodal degrees in a graph when we use statistical models to study tendencies toward higher-order network properties (such as reciprocity). We return to this idea in our discussion of digraphs, below. Statistics for degrees (means, variances, and so forth) and statistical distributions and inference are discussed in more detail in Chapters 5 and 13.

4.2.3 Density of Graphs and Subgraphs

Degree is a concept that considers the number of lines incident with each node in a graph. We can also consider the number and proportion of lines in the graph as a whole. A graph can only have so many lines. The maximum possible number is determined by the number of nodes. Since there are g nodes in the graph, and we exclude loops, there are $\binom{g}{2} = g(g-1)/2$ possible unordered pairs of nodes, and thus $g(g-1)/2$ possible lines that could be present in the graph. This is the maximum number of lines that can be present in a graph.

Consider now what proportion of these lines are actually present. The *density* of a graph is the proportion of possible lines that are actually present in the graph. It is the ratio of the number of lines present, L, to the maximum possible. The density of a graph, which we denote by Δ, is calculated as:

$$\Delta = \frac{L}{g(g-1)/2} = \frac{2L}{g(g-1)}. \tag{4.3}$$

The density of a graph goes from 0, if there are no lines present ($L = 0$), to 1, if all possible lines are present ($L = g(g-1)/2$).

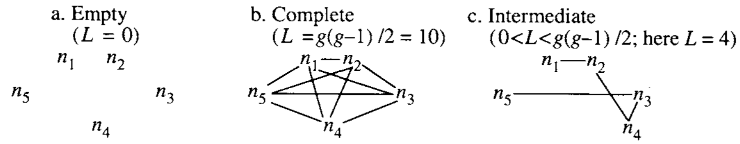


Fig. 4.4. Complete and empty graphs

If all lines are present, then all nodes are adjacent, and the graph is said to be *complete*. It is standard to denote a complete graph with g nodes as K_g. A complete graph contains all $g(g-1)/2$ possible lines, the density is equal to 1, and all nodal degrees are equal to $g - 1$.

An example of a complete graph in a social network would be a relation such as "communicates with," where all g actors communicated with all other actors.

There is a straightforward relationship between the density of a graph and the mean degree of the nodes in the graph. Noticing that the sum of the degrees is equal to $2L$ (since each line is counted twice, once for each of the two nodes incident with it — see equation (4.1)), we can combine equations (4.3) and (4.1) to get:

$$\Delta = \frac{\bar{d}}{(g-1)}. \tag{4.4}$$

In other words, the density of a graph is the average proportion of lines incident with nodes in the graph.

Figure 4.4 shows an example of an empty graph, a complete graph, and a graph with an intermediate number of lines ($L = 4$) for $g = 5$.

We can also define the density of a subgraph, which we will denote by Δ_s. The density of subgraph $\mathscr{G}_s$ is defined as the number of lines present in the subgraph, divided by the number of lines that could be present in the subgraph. We denote the number of nodes in subgraph $\mathscr{G}_s$ as g_s, and the number of lines in the subgraph as L_s. The possible number of lines in a subgraph is equal to $g_s(g_s - 1)/2$. We calculate the density of the subgraph as:

$$\Delta_s = \frac{2L_s}{g_s(g_s-1)}. \tag{4.5}$$

The density of a subgraph expresses the proportion of ties that are present among a subset of the actors in a network. This measure is used to evaluate the cohesiveness of subgroups (see Chapter 7) and to construct blockmodels and related simplified representations of networks (see Chapters 9 and 10).

We now turn to an example to demonstrate nodal degree and graph density.

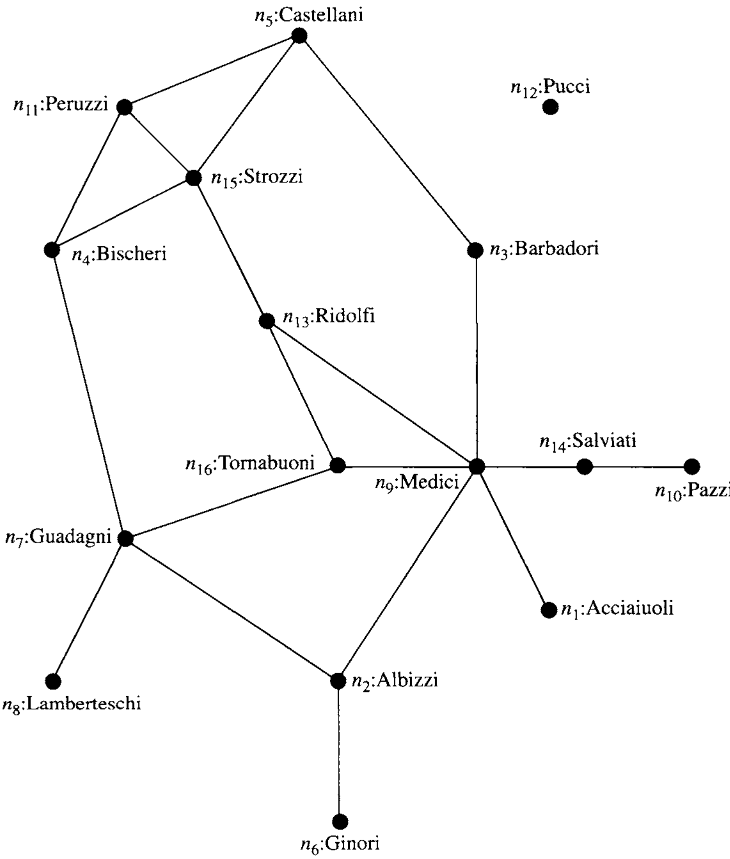
4.2.4 Example: Padgett's Florentine Families

Padgett's Florentine families network includes a set of sixteen Italian families in the early 15th century. The relation we consider here is marriage between pairs of families. Notice that the relation of marriage is nondirectional because the statement "a member of family i is married to a member of family j" is equivalent to the statement "a member of family j is married to a member of family i." In the graph, each family is represented as a node and the presence of a marriage between a pair of families is represented as a line. We label the sixteen nodes with the families' surnames.

The graph for this example is given in Figure 4.5. There are $g = 16$ nodes in this graph, and $L = 20$ lines between the pairs of nodes. Even with as few as sixteen actors and twenty ties, the graph looks rather complicated.

Let us now consider the nodal degrees and density of this social network. Notice that the Pucci family, n_{12}, is not related to any of the other families by the marriage relation. It is thus an isolate on the marriage relation. Note that the degrees sum to $2L = 40$. The mean nodal degree is $\bar{d} = (40/16) = 2.5$, the median is 3.0, and $S_D^2 = 1.46^2$. Thus, families have (on average) between two and three marriages to other families in this group. The Guadagni and Strozzi families, n_7 and n_{15}, have slightly more than average since their degrees are 4, and the Medici family, n_9, has quite a bit more than average with a degree of 6. The density of the graph is $20/120 = 0.167$.

Substantively, we know that much of the political posturing at this time in history centered on the Medici and the Strozzi families. Figure 4.5 indicates that these families seem to be important ones with respect to marital alliances. That is, these families have large degrees, and thus many marriages with the other families in this network. However, the Guadagni's, n_7, also have a large number of marriages.



node $= n_i$	*degree* $= d(n_i)$	*node* $= n_i$	*degree* $= d(n_i)$
n_1	1	n_9	6
n_2	3	n_{10}	1
n_3	2	n_{11}	3
n_4	3	n_{12}	0
n_5	3	n_{13}	3
n_6	1	n_{14}	2
n_7	4	n_{15}	4
n_8	1	n_{16}	3

Fig. 4.5. Graph and nodal degrees for Padgett's Florentine families, marriage relation

4.2.5 Walks, Trails, and Paths

In the previous section we considered the tie between a pair of nodes in terms of whether the nodes were adjacent or not. In this section we consider other ways in which two nodes can be linked by "indirect" routes that pass through the other nodes in the graph. We define and illustrate properties that are used to study the connectivity of graphs, to define the distance between pairs of nodes, and to identify nodes and lines that are critical for the connectivity of the graph.

These properties are not only important in themselves, but are also building blocks for later properties. In particular, *walks* and *paths* will allow us to calculate the distance between two nodes. We will define *walks,* their *inverses,* and measurement of the *lengths* of walks. We also describe special types of walks called *trails*, *paths*, *tours*, and *cycles*. Using the definition of paths, we define *geodesic distance*, *diameter*, and *eccentricity*.

In social network studies it is often important to know whether it is possible to reach some node n_i from another node n_j. If it is possible, it may also be interesting to know how many ways it can be done, and which of these ways is optimal with respect to one of several criteria. For example, we might wish to understand the communication of information among employees in an organization. An important consideration is whether information originating with one employee could eventually reach all other employees, and if so, how many lines it must traverse in order to get there. One might also consider whether there are multiple routes that a message might take to go from one employee to another, and whether some of these paths are more or less "efficient."

Walks in a Graph. A *walk* is a sequence of nodes and lines, starting and ending with nodes, in which each node is incident with the lines following and preceding it in the sequence. The listing of a walk, denoted by W, is an alternating sequence of incident nodes and lines beginning and ending with nodes. The beginning and ending nodes may be different. In addition, some nodes may be included more than once, and some lines may be included more than once. The *length* of a walk is the number of occurrences of lines in it. If a line is included more than once in the walk, it is counted each time it occurs.

Because (simple) graphs have at most one line between each pair of nodes, there is no ambiguity about which line is between any two nodes, and a walk may be described by just listing the nodes involved and excluding the lines. The starting node and the ending node of a walk

A walk would be $W = n_1\ l_2\ n_4\ l_3\ n_2\ l_3\ n_4$

A trail would be $W = n_4\ l_3\ n_2\ l_4\ n_3\ l_5\ n_4\ l_2\ n_1$

A path would be $W = n_1\ l_2\ n_4\ l_3\ n_2$

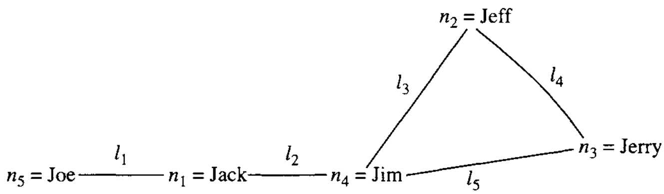


Fig. 4.6. Walks, trails, and paths in a graph

are the first and last nodes of W and are referred to as the *origin* and *terminus* of W. The *inverse* of a walk, denoted by W^{-1}, is the walk W listed in exactly the opposite order, using the same nodes and lines.

Figure 4.6 contains a hypothetical structure of communication ties among $g = 5$ employees. This structure might arise if actors Jeff, Jerry, and Jim (n_2, n_3, n_4) all communicate with each other and report through Jim (n_4) to Jack (n_1), who in turn reports to Joe (n_5). One possible walk through this network would be $W = n_1 l_2 n_4 l_3 n_2 l_3 n_4$. For example, perhaps Jack (n_1) passes a memo to Jim (n_4). Jim leaves the memo on Jeff's (n_2) desk. Jeff does not realize Jim was the one to leave the memo, and Jeff thinks he should bring the memo to Jim's attention, so he sends the memo back to Jim.

Notice several properties about this walk. First, not all nodes are involved: the message never reached Joe (n_5) or Jerry (n_3). Second, some nodes were used more than once: Jim (n_4) was included in the walk twice. Third, some lines were not used (that is, l_1, l_4, l_5), and some lines were used more than once (that is, l_3). The walk $W = n_1 l_2 n_4 l_3 n_2 l_3 n_4$ may be written more briefly as $W = n_1 n_4 n_2 n_4$. The origin and terminus in this walk are n_1 and n_4. The length of the walk is 3, since there are three lines: l_2, l_3, l_3. The length is 3 even though only two distinct lines are contained in this walk, because one of these lines is included twice. The inverse of the walk is $W^{-1} = n_4 n_2 n_4 n_1$.

A walk is the most general kind of sequence of adjacent nodes, since there are no restrictions on which nodes and lines may be included (aside from adjacency of nodes). Special kinds of walks, which we consider next, are more restrictive in that they require that nodes or lines be used no more than once.

○**Trails and Paths.** Trails and paths are walks with special characteristics. A *trail* is a walk in which all of the lines are distinct, though some node(s) may be included more than once. In the communications example, a trail means no communication tie is used more than once. The length of a trail is the number of lines in it.

A *path* is a walk in which all nodes and all lines are distinct. For example, a path through a communication network means no actor is informed more than once. The length of a path is the number of lines in it.

Notice that every path is a trail, and every trail is a walk. So any pair of nodes connected by a path is also connected by a trail and by a walk. Thus, a walk is the most general and a path is the least general kind of "route" through a graph. Since all paths are walks (but without repeating nodes or lines) a path is likely to be shorter compared to a walk or a trail. In a path in the communications network, no employee is informed more than once, and no pair of employees discusses the matter more than once. In applications to social networks, we will often focus on paths rather than walks.

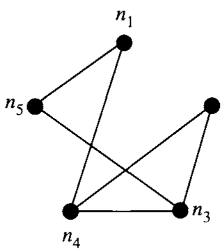
One of the trails in Figure 4.6 is $n_4n_2n_3n_4$ (no line is repeated). One of the paths is $n_1n_4n_2$ (no line or node is repeated).

There may be more than one path between a given pair of nodes. For example, in Figure 4.6 there are two paths between n_1 and n_2: $n_1n_4n_2$ and $n_1n_4n_3n_2$.

A very important property of a pair of nodes is whether there is a path between them, or not. If there is a path between nodes n_i and n_j, then n_i and n_j are said to be *reachable*. For example, if we consider a network of communications among people in which lines in a graph represent channels for transmission of messages between people, then if two actors are reachable, it is possible for a message to travel from one actor to the other by passing the message through intermediaries. If two actors are not reachable, then there is no path between them, and no way for a message to travel from one actor to the other.

○**Closed Walks, Tours, and Cycles.** Some walks begin and end at the same node. A walk that begins and ends at the same node is called a *closed walk*. In Figure 4.7, a closed walk is:

$$W = n_5n_1n_4n_3n_2n_4n_1n_5.$$



Tour $n_3\ n_2\ n_4\ n_3\ n_5\ n_1\ n_4\ n_3$

Cycles $n_5\ n_1\ n_4\ n_3\ n_5$
$n_2\ n_3\ n_4\ n_2$
$n_2\ n_4\ n_1\ n_5\ n_3\ n_2$

Closed walk $n_5\ n_1\ n_4\ n_3\ n_2\ n_4\ n_1\ n_5$

Fig. 4.7. Closed walks and cycles in a graph

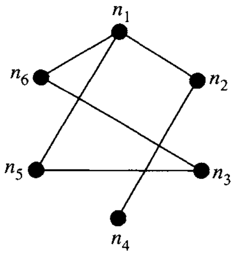
A *cycle* is closed walk of at least three nodes in which all lines are distinct, and all nodes except the beginning and ending node are distinct. A graph that contains no cycles is called *acyclic*. The graph in Figure 4.6 is not acyclic, since there is the cycle $W = n_4 n_2 n_3 n_4$.

A *tour* is a closed walk in which each line in the graph is used at least once. A tour in Figure 4.7 is:

$$W = n_3 n_2 n_4 n_3 n_5 n_1 n_4 n_3.$$

Cycles are important in the study of balance and clusterability in signed graphs (a topic we return to later in this chapter, and discuss in detail in Chapter 6).

Other special closed walks are those that include each and every node, or include each and every line. *Eulerian* trails are special closed trails that include every line exactly once (see Biggs, Lloyd, and Wilson 1976). Analogous closed walks can be defined in which each node is included exactly once. A cycle is labeled *Hamiltonian* if every node in the graph is included exactly once.



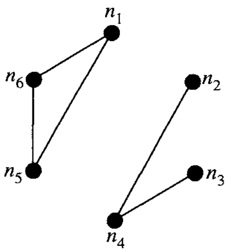


Fig. 4.8. A connected graph and a graph with components

4.2.6 Connected Graphs and Components

An important property of a graph is whether or not it is connected. A graph is *connected* if there is a path between every pair of nodes in the graph. That is, in a connected graph, all pairs of nodes are reachable. If a graph is not connected, then it is *disconnected.* Consider the example of communications among employees in an organization. If the graph representing communications among the employees is connected, then messages can travel from any employee to each and every other employee through the pairwise communication channels. However, if the graph representing this network is disconnected, then some pair of people cannot send or receive messages from each other using the communication channels.

Components. The nodes in a disconnected graph may be partitioned into two or more subsets in which there are no paths between the nodes in different subsets. The connected subgraphs in a graph are called *components.* A component of a graph is a maximal connected subgraph. Remember that a maximal entity is one that cannot be made larger and still retain its property. That is, a component is a subgraph in which there is a path between all pairs of nodes in the subgraph (all pairs of nodes in a component are reachable), and (since it is maximal) there is no path between a node in the component and any node not in the component. One cannot add another node to the subgraph and still retain the connectedness. If there is only one component in a graph, the graph is connected. If there is more than one component, the graph is disconnected.

Consider Figures 4.8a and 4.8b. The graph in Figure 4.8a is connected since there is a path between each pair of nodes. However, the graph in Figure 4.8b is not connected, since there is no path between n_1 and n_2. The graph in Figure 4.8b is *disconnected*, since there are pairs of nodes that do not have a path between them. For the graph in Figure 4.8b, the nodes can be partitioned into subsets $\mathscr{N}_1 = \{n_1, n_6, n_5\}$, $\mathscr{N}_2 = \{n_2, n_3, n_4\}$. The subgraphs generated by the different sets, $\mathscr{N}_1$ and $\mathscr{N}_2$ are the *components* of $\mathscr{G}$. In Figure 4.8b, the graph has two components.

Note that Padgett's Florentine families' marriage ties produce a disconnected graph because the Pucci family, represented by node n_{12}, is an isolate (that is, $d(n_{12}) = 0$). The two components in this graph are the subgraphs generated by the subsets:

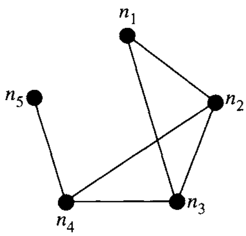
- $\mathscr{N}_1 = \{n_1, n_2, \ldots, n_{11}, n_{13}, \ldots, n_{16}\}$
- $\mathscr{N}_2 = \{n_{12}\}$

4.2.7 Geodesics, Distance, and Diameter

Now let us consider the paths between a pair of nodes. It is likely that there are several paths between a given pair of nodes, and that these paths differ in length. A shortest path between two nodes is referred to as a *geodesic*. If there is more than one shortest path between a pair of nodes, then there are two (or more) geodesics between the pair. The *geodesic distance* or simply the *distance* between two nodes is defined as the length of a geodesic between them. We will denote the geodesic distance between nodes n_i and n_j as $d(i, j)$. The distance between two nodes is the length of any shortest path between them. If there is no path between two nodes (that is, they are not reachable), then the distance between them is infinite (or undefined). If a graph is not connected, then the distance between at least one pair of nodes is infinite (because the distance between two nodes in different components is infinite). In a graph, a geodesic between n_i and n_j is also a geodesic between n_j and n_i. Thus the distance between n_i and n_j is equal to the distance between n_j and n_i; $d(i, j) = d(j, i)$.

Consider the graph in Figure 4.9. In this graph, the path $n_3n_4n_5$ is of length 2, since it contains two lines. This path is also a geodesic between n_3 and n_5; hence, $d(3, 5) = 2$ (the path $n_3n_2n_4n_5$ is of length 3 and is thus not a geodesic). Figure 4.9 also gives the geodesic distances between all pairs of nodes in this graph.

Geodesic distances



$d(1,2) = 1$
$d(1,3) = 1$
$d(1,4) = 2$
$d(1,5) = 3$
$d(2,3) = 1$
$d(2,4) = 1$
$d(2,5) = 2$
$d(3,4) = 1$
$d(3,5) = 2$
$d(4,5) = 1$

Diameter of graph = max $d(i, j) = d(1,5) = 3$

Fig. 4.9 Graph showing geodesics and diameter

Distances are quite important in social network analyses. They quantify how far apart each pair of nodes is, and are used in two of the centrality measures (discussed in Chapter 5) and are an important consideration for constructing some kinds of cohesive subgroups (discussed in Chapter 7).

○**Eccentricity of a Node.** Consider the geodesic distances between a given node and the other $g - 1$ nodes in a connected graph. The *eccentricity* or *association number* of a node is the largest geodesic distance between that node and any other node (Harary and Norman 1953; Harary 1969). Formally, the eccentricity of node n_i in a connected graph is equal to the maximum $d(i, j)$, for all j, (or $\max_j d(i, j)$). The eccentricity of a node can range from a minimum of 1 (if a node is adjacent to all other nodes in the graph) to a maximum of $g - 1$. It summarizes how far a node is from the node most distant from it in the graph. Several measures of centrality, such as the center and the centroid of a graph, are based on the eccentricity of the nodes. We discuss these in more detail in Chapter 5.

Diameter of a Graph. Consider the largest geodesic distance between any pair of nodes in a graph, that is, the largest eccentricity of any node. The *diameter* of a connected graph is the length of the largest geodesic between any pair of nodes (equivalently, the largest nodal eccentricity). Formally, the diameter of a connected graph is equal to the maximum $d(i, j)$, for all i and j (or $\max_i \max_j d(i, j)$). The diameter of a graph can range from a minimum of 1 (if the graph is complete) to a maximum of $g - 1$. If a graph is not connected, its diameter is infinite (or undefined) since the geodesic distance between one or more pairs of nodes in a disconnected graph is infinite.

Returning to the example in Figure 4.9 we see that the largest geodesic between any pair of nodes is 3 (between nodes n_1 and n_5). Thus the diameter of this graph is equal to 3.

The diameter of a graph is important because it quantifies how far apart the farthest two nodes in the graph are. Consider a communications network in which the ties are the transmission of messages. Focus on messages sent between all pairs of actors. Then, assuming messages always take the shortest routes (that is, via geodesics), we are guaranteed that a message can travel from any actor to any other actor, over a path of length no greater than the diameter of the graph.

Diameter of a Subgraph. We can also find the diameter of a subgraph. Consider a (node-generated) subgraph with node set $\mathcal{N}_s$ and line set $\mathcal{L}_s$ containing all lines from $\mathcal{L}$ between pairs of nodes in $\mathcal{N}_s$. The distance between a pair of nodes within the subgraph is defined for paths containing nodes from $\mathcal{N}_s$ and lines from $\mathcal{L}_s$. The distance between nodes n_i and n_j in the subgraph is the length of the shortest path between the nodes *within* the subgraph. Any path, and thus any geodesic, including nodes (and thus lines) outside the subgraph, is not considered. The *diameter of a subgraph* is the length of the largest geodesic within the subgraph.

4.2.8 Connectivity of Graphs

We now use the ideas of reachability between pairs of nodes, the concept of a connected graph, and components in a disconnected graph to define nodes and lines that are critical for the connectivity of a graph. We also present measures of how connected a graph is as a whole. The connectivity of a graph is a function of whether a graph remains connected when nodes and/or lines are deleted. We discuss each of these in turn.

Cutpoints. A node, n_i, is a *cutpoint* if the number of components in the graph that contains n_i is fewer than the number of components

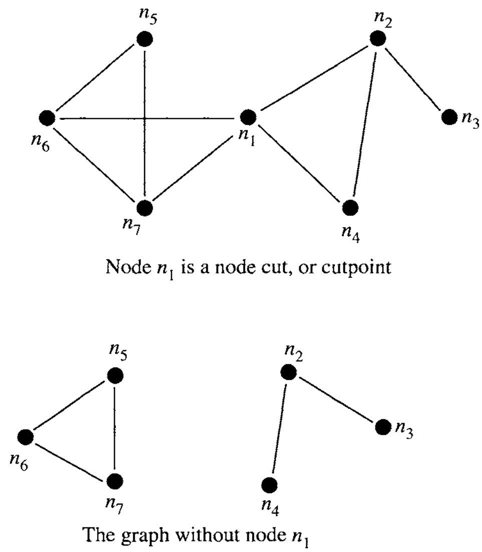


Fig. 4.10. Example of a cutpoint in a graph

in the subgraph that results from deleting n_i from the graph. That is, consider graph $\mathscr{G}$ with node set $\mathscr{N}$ which includes node n_i, and the subgraph $\mathscr{G}_s$ with node set $\mathscr{N}_s = \mathscr{N} - n_i$ that results from dropping n_i and all of its incident lines from graph $\mathscr{G}$. Node n_i is a cutpoint if the number of components in $\mathscr{G}$ is less than the number of components in $\mathscr{G}_s$.

For example, n_1 in Figure 4.10 is a cutpoint. This graph has one component, but if n_1 is removed, the graph has two components. In a communications network, an actor who is a cutpoint is critical, in the sense that if that actor is removed from the network, the remaining network has two subsets of actors, between whom no communication can travel.

The concept of a cutpoint can be extended from a single node to a *set* of nodes necessary to keep the graph connected. If a set of nodes is necessary to maintain the connectedness of a graph, these nodes are referred to as a *cutset*. If the set is of size k, then it is called a k-node cut. A cutpoint is a 1-node cut. If a set of nodes is a cutset, then the number of components in the graph that contains the set of nodes is fewer than

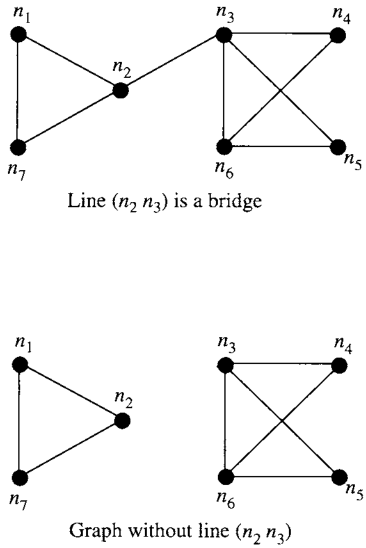


Fig. 4.11. Example of a bridge in a graph

the number of components in the subgraph that results from deleting the set of nodes from the graph.

Bridges. A notion analogous to that of cutpoint exists for lines. A bridge is a line that is critical to the connectedness of the graph. A *bridge* is a line such that the graph containing the line has fewer components than the subgraph that is obtained after the line is removed (nodes incident with the line remain in the subgraph). The removal of a bridge leaves more components than when the bridge is included. If line l_k is a bridge, then the graph $\mathscr{G}$ with line set $\mathscr{L}$ including l_k has fewer components than the subgraph $\mathscr{G}_s$ with line set $\mathscr{L} - l_k$, the graph obtained by deleting line l_k.

The line (n_2, n_3) in Figure 4.11 is a bridge. If the line (n_2, n_3) is removed from the graph, there is no path between nodes n_1 and n_5 (for example) and the graph becomes disconnected. In Figure 4.11, if the line (n_2, n_3) were nonexistent, nodes n_1, n_2, and n_7 would not be reachable from nodes n_3, n_4, n_5, and n_6.

Similarly, an *l-line cut* is a set of l lines that, if deleted, disconnects the graph. A bridge is a 1-line cut. In graphs representing social networks, a bridge is a critical tie, or a critical interaction between two actors.

Example. For the marriage relation for Padgett's Florentine families, the Medici family, n_9, is a cutpoint. With all sixteen nodes, the graph has two components. Without n_9, there are now two more components, giving four in total. If Family Medici is removed, n_1, the Acciaiuoli family becomes an isolate, and the Salviati and Pazzi families, n_{10} and n_{14}, are not reachable from the other families. There are other cutpoints in the graph. The marriage between the Salviati and Medici families, represented by the line (n_9, n_{14}), is a bridge (since its removal isolates the Salviati and Pazzi families).

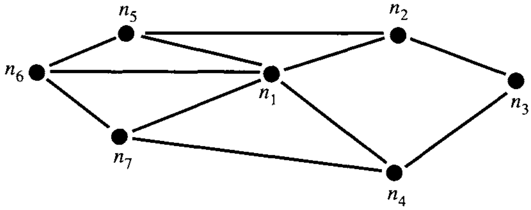
One can consider the extent of connectivity in a graph in terms of the number of nodes or the number of lines that must be removed in order to leave the graph disconnected. The connectivity of a graph is one measure of its "cohesiveness" or robustness.

⊗**Node- and Line-Connectivity.** One way to measure the cohesiveness of a graph is by its connectivity. A graph is cohesive if, for example, there are relatively frequent lines, many nodes with relatively large degrees, or relatively short or numerous paths between pairs of nodes. Cohesive graphs have many short geodesics, and small diameters, relative to their sizes. If a graph is not cohesive then it is "vulnerable" to the removal of a few nodes or lines. That is, a vulnerable graph is more likely to become disconnected if a few nodes or lines are removed.

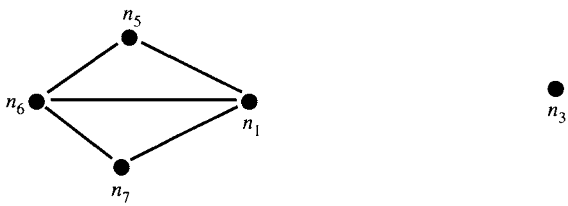
We can use the notions of a cutset and a line cut to define two measures of the connectivity of a graph. One measure describes the connectivity of the graph based on the removal of nodes, and the other describes the connectivity of the graph based on the removal of lines (Harary 1969).

The *point-connectivity* or *node-connectivity* of a graph, $\kappa(\mathscr{G})$, is the minimum number κ for which the graph has a κ-node cut. It is the minimum number of nodes that must be removed to make the graph disconnected, or to leave a trivial graph (Harary 1969, page 43). If the graph is disconnected, then $\kappa = 0$, since no node must be removed. If the graph contains a cutpoint, then $\kappa = 1$ since the removal of the single node leaves the graph disconnected. If a graph contains no node whose removal would disconnect the graph, but it contains a pair of nodes whose removal together would disconnect the graph, then $\kappa = 2$, since two is the minimum number of nodes that must be removed to make the graph disconnected. Thus, higher values of κ indicate higher levels of connectivity of the graph.

An example of a 2-node cut is given in Figure 4.12. The 2-node cut consists of n_2 and n_4, because without them n_3 would not be connected



n_2 and n_4 comprise a 2-node cut



The graph without n_2 and n_4

Fig. 4.12. Connectivity in a graph

to the remainder of the graph. In Figure 4.12, $\kappa(\mathscr{G}) = 2$. The graph may be disconnected if $\kappa \geq 2$ nodes are removed, but $\kappa = 2$ is the minimum. That is, the removal of any single node ($\kappa = 1$) would not result in a disconnected graph. In Figure 4.10, $\kappa(\mathscr{G}) = 1$, since there is a node whose removal disconnects the graph (n_1 is a cutpoint).

The value κ is the minimum number of nodes that must be removed to make the graph disconnected. Thus, removing any number of nodes less than κ does not make the graph disconnected. For any value k less than κ, the graph is said to be *k-node connected.*

A complete graph has no cutpoint; all nodes are adjacent to all others, so the removal of any one node would still leave the graph connected. In order to disconnect a complete graph, one would need to remove $g - 1$ nodes, resulting in a trivial graph ($g = 1$), so $\kappa(K_g) = g - 1$.

The *line-connectivity* or *edge-connectivity* of a graph, $\lambda(\mathscr{G})$, is the minimum number λ for which the graph has a λ-line cut. The value, λ, is the minimum number of lines that must be removed to disconnect the graph or leave a trivial graph (Harary 1969, page 43). In Figure 4.10, $\lambda(\mathscr{G}) = 1$, since line l_4 is a bridge. Removing more than one line may also destroy the graph's connectedness, but the minimum number of lines whose removal disconnects the graph is 1 (specifically line l_4). If $\lambda(\mathscr{G}) \geq l$, the graph is said to be *l-line connected*, since l is the minimum number of lines that must be removed to make the graph disconnected.

The larger the node-connectivity or the line-connectivity of a graph is, the less vulnerable the graph is to becoming disconnected. We will return to ideas of connectivity in Chapter 7 and discuss how these ideas can be used to define cohesive subgroups.

4.2.9 Isomorphic Graphs and Subgraphs

Two graphs, $\mathscr{G}$ and $\mathscr{G}^*$, are *isomorphic* if there is a one-to-one mapping from the nodes of $\mathscr{G}$ to the nodes of $\mathscr{G}^*$ that preserves the adjacency of nodes.

A one-to-one mapping means that each node in $\mathscr{G}$ is mapped to one (and only one) node in $\mathscr{G}^*$, and each node in $\mathscr{G}^*$ is mapped to one (and only one) node in $\mathscr{G}$. Let us denote nodes in $\mathscr{G}$ as $\mathscr{N} = \{n_1, n_2, \ldots, n_g\}$ and nodes in $\mathscr{G}^*$ as $\mathscr{N}^* = \{n_1^*, n_2^*, \ldots, n_g^*\}$. We will use the notation $\phi(n_i) = n_k^*$ to indicate that node n_i in $\mathscr{G}$ is mapped to node n_k^* in $\mathscr{G}^*$. The inverse of this mapping, ϕ^{-1}, is the mapping that maps node n_k^* in $\mathscr{G}^*$ to node n_i in $\mathscr{G}$; $\phi^{-1}(n_k^*) = n_i$. Since the mapping is a one-to-one mapping, $\phi(n_i) = n_k^*$ if and only if $\phi^{-1}(n_k^*) = n_i$.

The mapping preserves adjacency if nodes that are adjacent in $\mathscr{G}$ are mapped to nodes that are adjacent in $\mathscr{G}^*$, and nodes that are not adjacent in $\mathscr{G}$ are mapped to nodes that are not adjacent in $\mathscr{G}^*$, and vice versa. Formally, two graphs are isomorphic if for all $n_i, n_j \in \mathscr{N}$ and $n_k^*, n_l^* \in \mathscr{N}^*$ there exists a one-to-one mapping, $\phi(n_i) = n_k^*$ and $\phi(n_j) = n_l^*$ such that $l_m = (n_i, n_j) \in \mathscr{L}$ if and only if $l_o = (n_k^*, n_l^*) \in \mathscr{L}^*$. If two nodes are adjacent in one graph, then the nodes they are mapped to must also be adjacent in the isomorphic graph.

Consider the two graphs in Figure 4.13. Each graph has $g = 6$ nodes and $L = 6$ lines, and the nodes in each graph are labeled. A *labeled graph* is a graph in which the nodes have names or labels attached to them. The labels may be the names of the actors represented by the nodes, or they may be numbers or letters distinguishing the nodes. Isomorphic graphs are indistinguishable except for the labels on the nodes. For example, Figure 4.13a contains a graph $\mathscr{G}^*$ that is isomorphic to that in Figure 4.13b, $\mathscr{G}$.

Isomorphisms between graphs are important because if two graphs are isomorphic, then they are identical on all graph theoretic properties. For

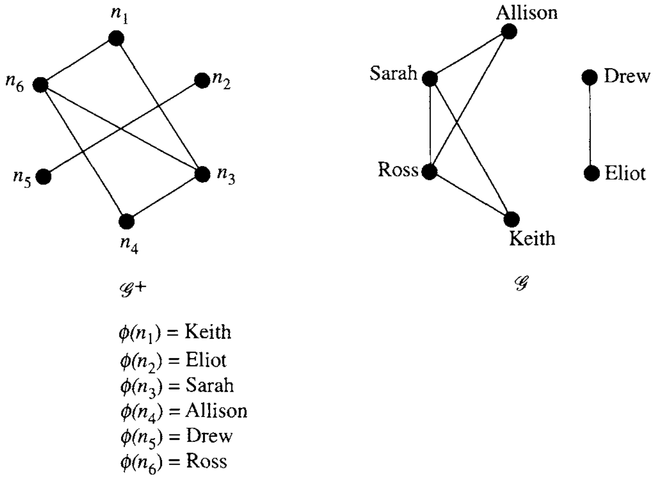


Fig. 4.13. Isomorphic graphs

example, two isomorphic graphs have the same number of nodes, the same number of lines, the same density, the same diameter, and so on. Thus, if we know that a particular graph theoretic property holds for graph $\mathscr{G}$ then we know that the property holds for any graph $\mathscr{G}^*$ that is isomorphic to $\mathscr{G}$.

It is also important to consider the nodes in isomorphic graphs. If two graphs $\mathscr{G}$ and $\mathscr{G}^*$ are isomorphic, and n_i in graph $\mathscr{G}$ is mapped to node n_k^* in $\mathscr{G}^*$, $(\phi(n_i) = n_k^*)$ then n_i and n_k^* are identical with respect to all graph theoretic properties (they have the same nodal degree, the same eccentricity, and so on). This property is quite important in defining some kinds of positional equivalences that we discuss in Chapter 12.

We can also consider isomorphic subgraphs. Two subgraphs, $\mathscr{G}_s$ and $\mathscr{G}_s^*$, are *isomorphic* if there is a one-to-one mapping from the nodes of $\mathscr{G}_s$ to the nodes of $\mathscr{G}_s^*$ that preserves the adjacency of nodes (as defined above). Subgraphs that are isomorphic belong to the same *isomorphism class*. Studies of dyads (Chapter 14) and triads (Chapter 15) rely on the isomorphism of very small subgraphs.

We now consider graphs with special properties.

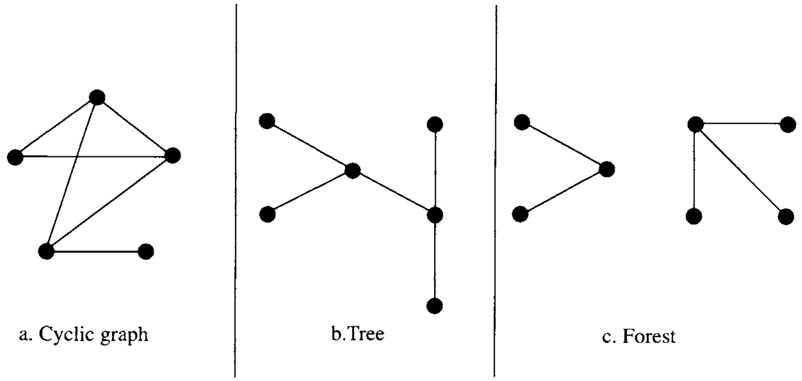


Fig. 4.14. Cyclic and acyclic graphs

4.2.10 ○Special Kinds of Graphs

Complement. The *complement*, $\overline{\mathscr{G}}$, of a graph, $\mathscr{G}$, has the same set of nodes as $\mathscr{G}$, a line is present between an unordered pair of nodes in $\overline{\mathscr{G}}$ if the unordered pair is *not* in the set of lines in $\mathscr{G}$, and a line is not present in $\overline{\mathscr{G}}$ if it is present in $\mathscr{G}$. In other words, if nodes n_i and n_j are adjacent in $\mathscr{G}$, then n_i and n_j are not adjacent in $\overline{\mathscr{G}}$, and if nodes n_i and n_j are not adjacent in $\mathscr{G}$, then n_i and n_j are adjacent in $\overline{\mathscr{G}}$. The line sets for these two graphs have no intersection at all, and their union is the set of all possible lines (all unordered pairs of nodes).

Trees. A graph that is connected and is acyclic (contains no cycles) is called a *tree*. In some ways trees are rather simple graphs, since they contain the minimum number of lines necessary to be connected, and they do not contain cycles. Several characteristics of trees are particularly important. First, trees are minimally connected graphs since every line in the graph is a bridge (or line cut). The removal of any one line causes the graph to be disconnected. Second, the number of lines in a tree equals the number of nodes minus one ($L = g - 1$). Adding another line adds a cycle to the graph, and hence the graph is no longer a tree. Third, there is only one path between any two nodes in a tree. If this is not true, the graph contains a cycle, which by definition a tree does not contain.

A graph that is disconnected (has more than one component) and contains no cycles is called a *forest*. In a forest, each component is a tree.

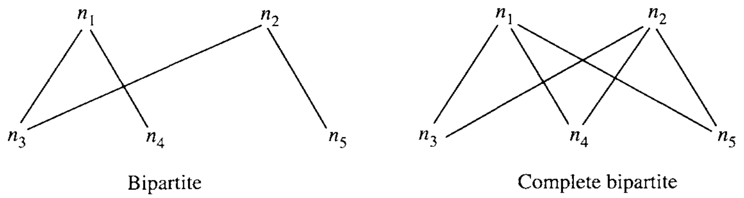


Fig. 4.15. Bipartite graphs

In general, the number of lines in a tree or forest equals the number of nodes minus the number of components of the graph. So, L equals g minus the number of components of $\mathscr{G}$. For a tree $L = g - 1$ since the number of components for a tree is 1.

The graph in Figure 4.14b is a tree. It is easy to verify that each pair of nodes is connected via some path, and the graph is acyclic. The graph in Figure 4.14a is not a tree, because it contains a cycle. The graph in Figure 4.14c is a forest. In the forest in Figure 4.14c, $L = 5$, or g minus 2 components.

Bipartite Graphs. If the nodes in a graph can be partitioned into two subsets, $\mathscr{N}_1$ and $\mathscr{N}_2$, so that every line in $\mathscr{L}$ is an unordered pair of nodes in which one node is in $\mathscr{N}_1$ and the other node is in $\mathscr{N}_2$, then the graph is *bipartite*. In a bipartite graph there are two subsets of nodes and all lines are between nodes belonging to different subsets. Nodes in a given subset are adjacent to nodes from the other subset, but no node is adjacent to any node in its own subset.

A *complete bipartite* graph is a bipartite graph in which every node in $\mathscr{N}_1$ is adjacent to every node in $\mathscr{N}_2$. Complete bipartite graphs are usually denoted K_{g_1,g_2}, where g_1 is the number of nodes in $\mathscr{N}_1$, and g_2 is the number of nodes in $\mathscr{N}_2$.

An example of a bipartite graph and a complete bipartite graph is given in Figure 4.15. Nodes n_1 and n_2 belong to $\mathscr{N}_1 = \{n_1, n_2\}$ and nodes n_3, n_4, n_5 belong to $\mathscr{N}_2 = \{n_3, n_4, n_5\}$.

A two-mode network with two sets of actors and a relation linking actors in one set to actors in the second set can be represented as a bipartite graph. But a bipartite graph may also exist in a one-mode network. A graph of an exogamous marriage system is bipartite, if, for example, women from clan A take husbands from clan B, and men from clan B take wives from clan A. In that case, all marriages unite partners from different clans.

The partitioning of the nodes in a graph can be generalized from two subsets $\mathcal{N}_1$ and $\mathcal{N}_2$ to s subsets $\mathcal{N}_1, \mathcal{N}_2, \ldots, \mathcal{N}_s$. An *s-partite* graph is one in which there is a partitioning of the nodes into s subsets so that all lines are between a node in $\mathcal{N}_i$ and a node in $\mathcal{N}_j$, where $i \neq j$. All lines are between nodes in different subsets and no nodes in the same subset are adjacent.

The notion of a complete bipartite graph can also be extended to a *complete s-partite graph.* A graph is a complete s-partite graph if all pairs of nodes belonging to different subsets are adjacent. All possible between-subset lines are present, and there are no lines incident with two nodes belonging to the same subset (equivalently, no nodes in the same subset are adjacent).

An example of a network that might be described by a bipartite graph is the set of monetary donations transacted between corporations in a specific geographic area, and the non-profit organizations headquartered in this area. We initially place all firms, both corporations and non-profit organizations, into a single actor set, $\mathcal{N}$. We then measure the flows of donations among these firms. Since the non-profit organizations usually have limited cash resources and thus can not support themselves financially, they must rely on the corporations for donations. We find that the only lines in this graph connect corporations to non-profit organizations. Thus, we have a bipartite graph, with the corporations residing in set $\mathcal{N}_1$ and non-profit organizations in set $\mathcal{N}_2$.

Thus far, we have focused our discussion on *graphs,* where a line between nodes is either present or absent. As we have emphasized before, graphs are useful for representing nondirectional relations. In the next section we discuss directed graphs, which are used for representing directional relations.

4.3 Directed Graphs

Many relations are *directional.* A relation is directional if the ties are oriented from one actor to another. The import/export of goods between nations is an example of a directional relation. Clearly goods go from one nation to another; one nation is the source and the other is the destination of the goods. In a social network representing trade among nations, the ties are directional and the graph representing such ties must be directed. Choices of friendships among children are another example of a directional relation. The claim of friendship is directed from one child to another child. Child i may choose child j as a friend, but that does not necessarily imply child j chooses child i as a friend.

In this section we define a directed graph and describe those definitions and concepts for directed graphs that are most useful for social network analysis. We refer the reader to Hage and Harary (1983), Harary, Norman, and Cartwright (1965), or other graph theory reference books for further discussion of directed graphs.

A directional relation can be represented by a *directed graph*, or *digraph* for short. A digraph consists of a set of nodes representing the actors in a network, and a set of arcs directed between pairs of nodes representing directed ties between actors. The difference between a graph and a directed graph is that in a directed graph the direction of the lines is specified. Directed ties between the pairs of actors are represented as lines in which the orientation of the relation is specified. These oriented lines are called *arcs*.

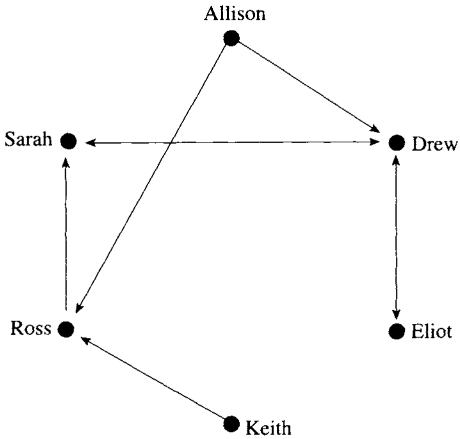
A *directed graph*, or *digraph*, $\mathscr{G}_d(\mathscr{N}, \mathscr{L})$, consists of two sets of information: a set of nodes $\mathscr{N} = \{n_1, n_2, \ldots, n_g\}$, and a set of arcs, $\mathscr{L} = \{l_1, l_2, \ldots, l_L\}$. Each arc is an *ordered pair* of distinct nodes, $l_k =< n_i, n_j >$. The arc $< n_i, n_j >$ is directed from n_i (the origin or sender) to n_j (the terminus or receiver). The difference between an arc (in a digraph) and a line (in a graph) is that an arc is an *ordered pair* of nodes (to reflect the direction of the tie between the two nodes) whereas a line is an unordered pair of nodes (it simply records the presence of a tie between two nodes).

We let L be the number of arcs in $\mathscr{L}$. Since each arc is an ordered pair of nodes, there are $g(g-1)$ possible arcs in $\mathscr{L}$.

As in a graph, a node is *incident* with an arc if the node is in the ordered pair of nodes defining the arc. For example, both nodes n_i and n_j are incident with the arc $l_k =< n_i, n_j >$. However, in a digraph, since an arc is an ordered pair of nodes, we can distinguish the first from the second node in the pair. Thus, the concept of adjacency of pairs of nodes in a digraph is somewhat more complicated than adjacency of pairs of nodes in a graph. We must consider whether a given node is first (sender) or second (receiver) in the ordered pair defining the arc. Specifically, node n_i is *adjacent to* node n_j if $< n_i, n_j > \in \mathscr{L}$, and node n_j is *adjacent from* node n_i if $< n_i, n_j > \in \mathscr{L}$.

When a digraph is presented as a diagram the nodes are represented as points and the arcs are represented as directed arrows. The arc $< n_i, n_j >$ is represented by an arrow from the point representing n_i to the point representing n_j. For example, if actor i nominates actor j as a friend,

	Actor	Likes at beginning of year
n_1	Allison	Drew Ross
n_2	Drew	Eliot Sarah
n_3	Eliot	Drew
n_4	Keith	Ross
n_5	Ross	Sarah
n_6	Sarah	Drew



$l_1 = <n_1, n_2>$ $\quad l_5 = <n_3, n_2>$

$l_2 = <n_1, n_5>$ $\quad l_6 = <n_4, n_5>$

$l_3 = <n_2, n_3>$ $\quad l_7 = <n_5, n_6>$

$l_4 = <n_2, n_6>$ $\quad l_8 = <n_6, n_2>$

Fig. 4.16. Friendship at the beginning of the year for six children

there would be an arc originating at i and terminating at j. If actor j returned the friendship choice, there would be another arc, this one originating at j and terminating at i.

To illustrate a directed graph let us consider the choices of friendship among our six children at the beginning of the year. These choices are represented in the directed graph in Figure 4.16. The $g = 6$ nodes represent the six children, and the arcs represent friendship nominations. So, there is an arc from one node to another if the child represented by the first node chose the child represented by the second node as a friend. For example Ross, n_5, chose Sarah, n_6, as a friend, so the arc $< n_5, n_6 >$ is included in the graph.

Many concepts for graphs (such as subgraph) presented and defined earlier in this chapter are immediately applicable to directed graphs, and thus do not require special discussion. However, some concepts, such as isomorphism classes for dyads and triads, nodal degree, walks, and paths are somewhat different in directed graphs, and thus need special discussion. We now turn to these digraph topics.

4.3.1 Subgraphs – Dyads

One of the most important subgraphs in a digraph is the dyad, consisting of two nodes and the possible arcs between them. Since there may or may not be an arc in either direction for a pair of nodes, n_i and n_j, there are four possible states for each dyad. However, there are only three isomorphism classes (all dyads are identical to one of these three types).

The first isomorphism class of a dyad is the *null* dyad. Null dyads have no arcs, in either direction, between the two nodes. The dyad for nodes n_i and n_j is null if neither of the arcs $< n_i, n_j >$ *nor* $< n_j, n_i >$ is contained in the set of arcs, $\mathscr{L}$. The second isomorphism class is called *asymmetric*. An asymmetric dyad has an arc between the two nodes going in one direction or the other, but not both. The dyad for nodes n_i and n_j is asymmetric if either one of the arcs $< n_i, n_j >$ *or* $< n_j, n_i >$, but not both, is contained in the set of arcs, $\mathscr{L}$. Thus, there are two possible asymmetric dyads, but they are isomorphic. The third isomorphism class is called a *mutual* or *reciprocal* dyad. Mutual dyads have two arcs between the nodes, one going in one direction and the other going in the opposite direction. The dyad for nodes n_i and n_j is mutual if *both* arcs $< n_i, n_j >$ *and* $< n_j, n_i >$ are contained in the set of arcs, $\mathscr{L}$. Thus the three isomorphism classes for dyads are: null, asymmetric, and mutual.

If the directed graph represents the friendship relation, a null dyad is one in which neither person chooses the other. The asymmetric dyad occurs when one person chooses the other, without the choice being reciprocated. In a mutual dyad both actors in the pair choose each other as friends.

Figure 4.17 shows the dyads for the example of friendships at the beginning of the year among the six children (presented in Figure 4.16). Since there are $g = 6$ children, there are $6(5 - 1)/2 = 15$ dyads to consider. Figure 4.17 shows the state of each of these 15 dyads.

The arc with a double-headed arrow between n_2 and n_3 indicates a mutual dyad. Asymmetric dyads are represented by one-way arcs, such

n_1	$\longrightarrow$	n_2	(asymmetric)
n_1		n_3	(null)
n_1		n_4	(null)
n_1	$\longrightarrow$	n_5	(asymmetric)
n_1		n_6	(null)
n_2	$\longleftrightarrow$	n_3	(mutual)
n_2		n_4	(null)
n_2		n_5	(null)
n_2	$\longleftrightarrow$	n_6	(mutual)
n_3		n_4	(null)
n_3		n_5	(null)
n_3		n_6	(null)
n_4	$\longrightarrow$	n_5	(asymmetric)
n_4		n_6	(null)
n_5	$\longrightarrow$	n_6	(asymmetric)

Fig. 4.17. Dyads from the graph of friendship among six children at the beginning of the year

as from n_1 to n_2. The dyad involving Allison, n_1, and Keith, n_4, is a null dyad, since neither arc is present.

The kinds of dyads that arise in a directed graph are quite interesting and important for describing a social network. Tendencies for reciprocity (mutuality) and/or asymmetry in a digraph are often summarized by counting the number of dyads in each of the three isomorphism classes. Chapter 13 discusses these ideas and presents some models for dyads.

One could study subgraphs of any size for a digraph. Dyads are clearly subgraphs of size two. Triads, subgraphs of size three, are important for studying ideas such as balance, clusterability, and transitivity (which we describe in detail in Chapter 6). Cohesive subgroups are also studied by focusing on subgroups (see Chapter 7).

We now discuss how several of the concepts for graphs are applied to directed graphs. We will focus on the most important directed graph concepts including the nodal degrees, walks, paths, reachability, and connectivity.

4.3.2 Nodal Indegree and Outdegree

In a graph, the degree of a node is the number of nodes adjacent to it (equivalently, the number of lines incident with it). In a digraph, a node can be either *adjacent to*, or *adjacent from* another node, depending on the "direction" of the arc. Thus, it is necessary to consider these cases separately. One quantifies the tendency of actors to make "choices"; the other quantifies the tendency to receive "choices."

The *indegree* of a node, $d_I(n_i)$, is the number of nodes that are adjacent *to* n_i. The indegree of node n_i is equal to the number of arcs of the form $l_k =< n_j, n_i >$, for all $l_k \in \mathscr{L}$, and all $n_j \in \mathscr{N}$. Indegree is thus the number of arcs terminating at n_i.

The *outdegree* of a node, $d_O(n_i)$, is the number of nodes adjacent *from* n_i. The outdegree of node n_i is equal to the number of arcs of the form $l_k =< n_i, n_j >$, for all $l_k \in \mathscr{L}$, and all $n_j \in \mathscr{N}$. Outdegree is thus the number of arcs originating with node n_i.

The indegrees and outdegrees for each node may be obtained by considering the arcs in the digraph. Thus, the outdegrees for the six nodes, representing children, in Figure 4.16 are:

- $d_O(n_1) = 2$
- $d_O(n_2) = 2$
- $d_O(n_3) = 1$
- $d_O(n_4) = 1$
- $d_O(n_5) = 1$
- $d_O(n_6) = 1$

The indegrees are:

- $d_I(n_1) = 0$
- $d_I(n_2) = 3$
- $d_I(n_3) = 1$
- $d_I(n_4) = 0$
- $d_I(n_5) = 2$
- $d_I(n_6) = 2$

In social network applications, these degrees can be of great interest. The outdegrees are measures of *expansiveness* and the indegrees are measures of *receptivity*, or *popularity*. If we consider the sociometric relation of friendship, an actor with a large outdegree is one who nominates many others as friends. An actor with a small outdegree nominates fewer friends. An actor with a large indegree is one whom many others nominate as a friend, and an actor with a small indegree is chosen by few others. Outdegrees may be fixed by the data collection design, if, for example, a researcher collects data in which each respondent is instructed to "name your three closest friends." In such a setting, if all respondents in fact named three closest friends, then all outdegrees would equal 3.

Indegrees and outdegrees are useful measurements for many different types of networks and relations, although the terms "expansive" and "popular" may be somewhat inappropriate in some cases. For example, consider the countries trade network, and the relation "exports manufactured goods to" among countries. A country with high outdegree is a heavy exporter, and a country with high indegree is a heavy importer.

In many statistical models we might want to control for, or condition on, either the indegrees or the outdegrees of the nodes. For example, if we are studying the tendency for mutual choices within a network, we might control for the nodal outdegrees; that is, we would study the tendency for mutuality, given the propensity of our actors to make choices. Such statistical conditioning is used in Chapters 13–16.

It is often useful to summarize the indegrees and/or the outdegrees of all the actors in the network using the mean indegree or the mean outdegree. As we will see, these two numbers are equal, since they are considering the same set of arcs, but from different "directions." We will denote the mean indegree as $\bar{d}_I$, and the mean outdegree as $\bar{d}_O$. These are calculated as:

$$\begin{aligned} \bar{d}_I &= \frac{\sum_{i=1}^{g} d_I(n_i)}{g} \\ \bar{d}_O &= \frac{\sum_{i=1}^{g} d_O(n_i)}{g}. \end{aligned} \tag{4.6}$$

Since the indegrees count arcs incident from the nodes, and the outdegree count arcs incident to the nodes, $\sum_{i=1}^{g} d_I(n_i) = \sum_{i=1}^{g} d_O(n_i) = L$, and thus we can see that $\bar{d}_I = \bar{d}_O$ and equations (4.6) simplify to:

$$\bar{d}_I = \bar{d}_O = \frac{L}{g}. \tag{4.7}$$

One might also be interested in the variability of the nodal indegrees and outdegrees. Unlike the mean indegree and the mean outdegree, the variance of the indegrees is not necessarily the same as the variance of the outdegrees. For example, consider a sociometric question in which each person is asked to name her three closest friends. If all people in fact make three nominations, then there is no variance in the outdegrees (all $d_O(n_i) = 3$). However, it is likely that people will receive different numbers of "choices"; thus, there will be variability in the indegrees (the $d_I(n_i)$'s will differ from each other). The variance of the indegrees, which we denote by $S^2_{D_I}$, is calculated as:

$$S_{D_I}^2 = \frac{\sum_{i=1}^{g}(d_I(n_i) - \bar{d}_I)^2}{g}. \quad (4.8)$$

Similarly, the variance of the outdegrees, which we denote by $S_{D_O}^2$, is calculated as:

$$S_{D_O}^2 = \frac{\sum_{i=1}^{g}(d_O(n_i) - \bar{d}_O)^2}{g}. \quad (4.9)$$

Both of these measures quantify how unequal the actors in a network are with respect to initiating or receiving ties. These measures are simple statistics for summarizing how "centralized" a network is. We return to this idea in Chapter 5.

Types of Nodes in a Directed Graph. The indegrees and outdegrees of the nodes in a directed graph can be used to distinguish four different kinds of nodes based on the possible ways that arcs can be incident with the node. Recall that the indegree of node n_i, denoted by $d_I(n_i)$, is equal to the number of nodes adjacent to it, and the outdegree of node n_i, denoted by $d_O(n_i)$, is equal to the number of nodes adjacent from it. In terms of the indegree and outdegree there are four possible kinds of nodes: the node is an isolate, the node only has arcs originating from it, the node only has arcs terminating at it, or the node has arcs both to and from it. Graph theorists provide a vocabulary for labeling these four kinds of nodes (Harary, Norman, and Cartwright 1965, page 18; Hage and Harary 1983). According to this classification, a node is a(n):

- *Isolate* if $d_I(n_i) = d_O(n_i) = 0$,
- *Transmitter* if $d_I(n_i) = 0$ and $d_O(n_i) > 0$,
- *Receiver* if $d_I(n_i) > 0$ and $d_O(n_i) = 0$,
- *Carrier* or *ordinary* if $d_I(n_i) > 0$ and $d_O(n_i) > 0$

The distinction between a carrier and an ordinary node is that, although both kinds have both positive indegree and positive outdegree, a carrier has both indegree and outdegree precisely equal to 1, whereas an ordinary node has indegree and/or outdegree greater than 1.

Several authors have argued that this typology, or some variant of it, is useful for describing the "roles" or "positions" of actors in social networks (Burt 1976; Marsden 1989; Richards 1989a).

4.3.3 Density of a Directed Graph

The density of a directed graph is equal to the proportion of arcs present in the digraph. It is calculated as the number of arcs, L, divided by the possible number of arcs. Since an arc is an ordered pair of nodes, there are $g(g-1)$ possible arcs. The density, Δ, is:

$$\Delta = \frac{L}{g(g-1)}. \tag{4.10}$$

The density of a digraph is a fraction that goes from a minimum of 0, if no arcs are present, to a maximum of 1, if all arcs are present. If the density is equal to 1, then all dyads are mutual.

4.3.4 An Example

Now let us illustrate nodal indegree and outdegree, and the density of a directed graph on the example of friendships among Krackhardt's high-tech managers. Clearly a directed graph is the appropriate representation for these friendship choices, since each choice of friendship is directed from one manager to another (and is not necessarily reciprocated).

Table 4.1 presents the nodal indegrees and outdegrees, the mean and variance of the indegrees and outdegrees, and the density of the graph. From these results we see that there are no isolates in this network (there are no managers with both indegree and outdegree equal to 0). However, there are two managers (managers 7 and 9) who did not make any friendship nominations. The mean number of friendship choices made (and received) is equal to 4.86. The density of the relation is equal to 0.243.

4.3.5 Directed Walks, Paths, Semipaths

Walks and related concepts in graphs can also be defined for digraphs, but one must consider the direction of the arcs. We first define directed walks, directed paths, and semipaths for directed graphs and then define closed walks (cycles and semicycles) for directed graphs.

A *directed walk* is a sequence of alternating nodes and arcs so that each arc has its origin at the previous node and its terminus at the subsequent node. More simply, in a directed walk, all arcs are "pointing" in the same direction. The length of a directed walk is the number of instances of arcs in it (an arc is counted each time it occurs in the walk).

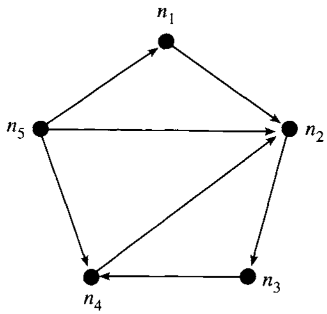
Table 4.1. *Nodal degree and density for friendships among Krackhardt's high-tech managers*

Manager	Indegree	Outdegree
1	8	5
2	10	3
3	5	2
4	5	6
5	6	7
6	2	6
7	3	0
8	5	1
9	6	0
10	1	7
11	6	13
12	8	4
13	1	2
14	5	2
15	4	8
16	4	2
17	6	18
18	4	1
19	5	9
20	3	2
21	5	4

$L = 102$
$g = 21$
$\bar{d}_I = \bar{d}_O = 102/21 = 4.86$
possible number of arcs: $21(20) = 420$
$\Delta = 102/420 = 0.243$
$S^2_{D_I} = 2.17^2$
$S^2_{D_O} = 4.37^2$

For example, consider the digraph in Figure 4.18. One directed walk in this figure is $W = n_5 n_1 n_2 n_3 n_4 n_2 n_3$.

Recall that a trail in a graph is a walk in which no line is included more than once. A *directed trail* in a digraph is a directed walk in which no arc is included more than once. Similarly, a *directed path* or simply a *path* in a digraph is a directed walk in which no node and no arc is included more than once. A path joining nodes n_i and n_j in a directed graph is a sequence of distinct nodes, where each arc has its origin at the previous node, and its terminus at the subsequent node. Thus, a path in a directed graph consists of arcs all "pointing" in the same direction. The length of a path is the number of arcs in it.



Directed walk	$n_5\ n_1\ n_2\ n_3\ n_4\ n_2\ n_3$
Directed path	$n_5\ n_4\ n_2\ n_3$
Semipath	$n_1\ n_2\ n_5\ n_4\ n_3$
Cycle	$n_2\ n_3\ n_4\ n_2$
Semicycle	$n_1\ n_2\ n_5\ n_1$

Fig. 4.18. Directed walks, paths, semipaths, and semicycles

Now, consider removing the restriction that all arcs "point" in the same direction. We will simply consider walks and paths in which the arc between previous and subsequent nodes in the sequence may go in either direction. A *semiwalk* joining nodes n_i and n_j is a sequence of nodes and arcs in which successive pairs of nodes are incident with an arc from the first to the second, or by an arc from the second to the first. That is, in a semiwalk, for all successive pairs of nodes, the arc between adjacent nodes may be either $< n_i, n_j >$ or $< n_j, n_i >$. In a semiwalk the direction of the arcs is irrelevant. The length of a semiwalk is the number of instances of arcs in it.

A *semipath* joining nodes n_i and n_j is a sequence of distinct nodes, where all successive pairs of nodes are connected by an arc from the first to the second, or by an arc from the second to the first for all successive pairs of nodes (Harary, Norman, and Cartwright 1965; Peay 1975). In a semipath the direction of the arcs is irrelevant. The length of a semipath is the number of arcs in it.

Note that every path is a semipath, but not every semipath is a path (see Harary, Norman, and Cartwright 1965, for more discussion).

Closed walks can also be defined for directed graphs. A *cycle* in a directed graph is a closed directed walk of at least three nodes in which all nodes except the first and last are distinct. A *semicycle* in a directed graph is a closed directed semiwalk of at least three nodes in which all nodes except the first and last are distinct. In a semicycle the arcs may go in either direction, whereas in a cycle the arcs must all "point" in the same direction. Semicycles are used to study structural balance and clusterability (see Chapter 6).

Figure 4.18 gives examples of a directed walk, a directed path, a semipath, a cycle, and a semicycle.

4.3.6 Reachability and Connectivity in Digraphs

Using the ideas of paths and semipaths, we can now define reachability and connectivity of pairs of nodes, and the connectedness of a directed graph.

Pairs of Nodes. In a graph a pair of nodes is reachable if there is a path between them. However, in order to define reachability in a directed graph, we must consider directed paths. Specifically, if there is a directed path from n_i to n_j, then node n_j is *reachable from* node n_i.

Consider now both paths and semipaths between pairs of nodes. We can define four different ways that two nodes can be connected by a path, or semipath (Harary, Norman, and Cartwright 1965; Frank 1971; Peay 1975, 1980). A pair of nodes, n_i, n_j, is:

(i) *Weakly connected* if they are joined by a *semipath*

(ii) *Unilaterally connected* if they are joined by a *path* from n_i to n_j, *or* a *path* from n_j to n_i

(iii) *Strongly connected* if there is a *path* from n_i to n_j, *and* a path from n_j to n_i; the path from n_i to n_j may contain different nodes and arcs than the path from n_j to n_i

(iv) *Recursively connected* if they are strongly connected, and the path from n_i to n_j uses the same nodes and arcs as the path from n_j to n_i, in reverse order

Notice that these forms of connectivity are increasingly strict, and that any strict form implies connectivity of any less strict form. For example, any two nodes that are recursively connected are also strongly connected, unilaterally connected, and weakly connected. Figure 4.19 illustrates these different kinds of connectivity. In each case nodes n_1 and n_4 in the graph demonstrate the different versions of connectivity.

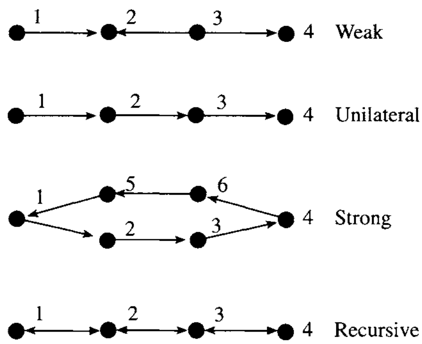


Fig. 4.19. Different kinds of connectivity in a directed graph

Digraph Connectedness. It is now possible to define four different kinds of connectivity for digraphs (Peay 1975, 1980). If a digraph is connected, then it is connected by one of these four kinds of connectivity; otherwise, it is not connected. Since there are four types of connectivity between pairs of nodes in a directed graph, there are four definitions of graph connectivity for a digraph. A directed graph is:

(i) *Weakly connected* if all pairs of nodes are weakly connected
(ii) *Unilaterally connected* if all pairs of nodes are unilaterally connected
(iii) *Strongly connected* if all pairs of nodes are strongly connected
(iv) *Recursively connected* if all pairs of nodes are recursively connected

In a weakly connected digraph, all pairs of nodes are connected by a semipath. In a unilaterally connected digraph, between each pair of nodes there is a directed path from one node to the other; in other words at least one node is reachable from the other in the pair. In a strongly connected digraph each node in each pair is reachable from the other; there is a directed path from each node to each other node. In a recursively connected digraph, each node, in each pair, is reachable from the other, and the directed paths contain the same nodes and arcs, but in reverse order. As with the definitions of connectivity for pairs of nodes, these are increasingly strict graph connectivity definitions.

From these definitions it should be clear that every strongly connected digraph is unilaterally connected, but the reverse is not true. When maximal subgraphs are derived from digraphs in which the actors are unilaterally, or strongly, connected, the subgraph is referred to as a *unilateral*, or *strong*, component in the digraph. These ideas are used to study cohesive subgroups in directed graphs (see Chapter 7).

4.3.7 Geodesics, Distance and Diameter

The (geodesic) distance between a pair of nodes in a graph is the length of a shortest path between the two nodes, and is the basis for defining the diameter of the graph. In a directed graph, the paths from node n_i to node n_j may be different from the paths from node n_j to node n_i (because paths in a directed graph consider the direction of the arcs). Thus, the definitions of distance and diameter in a directed graph are somewhat more complicated than in a graph.

Consider the paths from node n_i to node n_j. A geodesic from node n_i to node n_j is a shortest path *from* n_i to n_j. The distance from n_i to n_j, denoted by $d(i, j)$, is the length of a geodesic from n_i to n_j. It is important to note that since the paths from n_i are likely to be different from the paths from n_j to n_i (since paths require that all arcs are "pointing" in the same direction) the geodesics from n_i to n_j may be different from the geodesics from n_j to n_i. Thus, the distance, $d(i, j)$, from n_i to n_j may be different from the distance, $d(j, i)$, from n_j to n_i. For example, in Figure 4.18 $d(4, 2) = 1$ whereas $d(2, 4) = 2$. If there is no path from n_i to n_j (as might be the case when the graph is only weakly or unilaterally connected) then there is no geodesic from n_i to n_j, and the distance from n_i to n_j is undefined (or infinite).

Now, consider the diameter of a directed graph. As in a graph, the diameter of a directed graph is the length of the longest geodesic between any pair of nodes. This definition of geodesic is useful if there is a *path* from each node to each other node in the graph; that is, the graph is strongly connected or recursively connected. However, if the graph is only unilaterally or weakly connected, then, as noted above, some distances are undefined (or infinite). Thus, the diameter of a weakly or unilaterally connected directed graph is undefined.

4.3.8 ○Special Kinds of Directed Graphs

In this section we describe several kinds of digraphs with important properties. We begin by defining digraph complement and digraph converse.

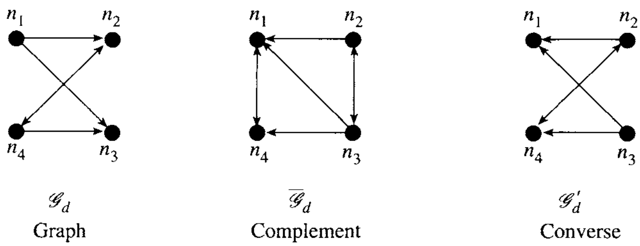


Fig. 4.20. Converse and complement of a directed graph

Complement and Converse of a Digraph. Now let us consider two kinds of digraphs that can be derived from a digraph. These derived digraphs can be used to represent the opposite and the negation of a relation.

The *complement*, $\overline{\mathscr{G}}_d$, of a directed graph, $\mathscr{G}_d$, has the same set of nodes as $\mathscr{G}_d$, but there is an arc present between an ordered pair of nodes in $\overline{\mathscr{G}}_d$ if the ordered pair is *not* in the set of arcs in $\mathscr{G}_d$, and an arc is not present in $\overline{\mathscr{G}}_d$ if it is present in $\mathscr{G}_d$. In other words, if the arc $< n_i, n_j >$ is in $\mathscr{G}_d$, then the arc $< n_i, n_j >$ is not in $\overline{\mathscr{G}}_d$, and if the arc $< n_i, n_j >$ is not in $\mathscr{G}_d$, then the arc $< n_i, n_j >$ is in $\overline{\mathscr{G}}_d$.

The *converse*, $\mathscr{G}'_d$, of a directed graph, $\mathscr{G}_d$, has the same set of nodes as $\mathscr{G}_d$, but the arc $< n_i, n_j >$ is in $\mathscr{G}'_d$ only if the arc $< n_j, n_i >$ is in $\mathscr{G}_d$ (Harary 1969). The converse, $\mathscr{G}'_d$, is obtained from $\mathscr{G}_d$ by reversing the direction of all arcs. The arcs in the converse connect the same pairs of nodes as the arcs in the digraph, but all arcs are reversed in direction. That is, an arc in the digraph from n_i to n_j becomes an arc in the converse from n_j to n_i. Figure 4.20 shows a directed graph, its converse, and its complement.

The converse of a directed graph might be helpful in thinking about relations that have "opposites." For example, the converse of a digraph representing a dominance relation (for example, n_i "wins over" n_j) would represent the submissive relation (n_j "loses to" n_i). On the other hand, the complement of a digraph might be used to represent the absence of a tie, or as *not* the relation. For example, in the digraph representing the relation of friendship the arc $< n_i, n_j >$ means i "chooses" j as a friend. In the digraph representing the complement of the relation of friendship, the arc $< n_i, n_j >$ means i "does not choose" j as a friend.

Tournaments. One other special type of a digraph is a *tournament*, which mathematically represents a set of actors competing in some event(s) and a relation indicating superior performances or "beats" in competition (see Moon 1968). If team n_i beats team n_j, an arc is directed from n_i toward n_j. Of particular interest are *round-robin* tournaments, where each team plays each other team exactly once. Such tournaments can be modeled as *round robin designs* (Kenny 1981; Kenny and LaVoie 1984; Wong 1982). These competitive records form a special type of digraph, because each pair of nodes is connected by exactly one arc. Methodology for such designs is related to the Bradley-Terry-Luce model for paired comparisons, which allows for statistical estimation of population propensities for dominance (Bradley and Terry 1952; Thurstone 1927; Coombs 1951; Mosteller 1951; Frank 1981; and David 1988).

4.3.9 Summary

Digraphs are the appropriate representation of social networks in which relations are dichotomous (ties are either present or absent) and directional. However, many relations are valued; that is, the ties indicate the strength or intensity of the tie between each pair of actors. Thus, we need to generalize both graphs and directed graphs so that we can represent the *strength* of ties between actors in a network. The graph for a valued relation must convey more information by representing the strength of an arc or a line. For example, observations of the number of interactions between pairs of people in a group require valued relations. Similarly, ratings of friendship in which people distinguish between "close personal friends," "friends," "acquaintances," and "strangers" must be represented by a graph in which the arcs also have a value indicating the strength of the tie. In the next sections we define and discuss signed graphs (in which the lines or arcs take on a positive or negative sign). In the section following that we discuss valued graphs (in which the lines or arcs can take on a values from the real numbers).

4.4 Signed Graphs and Signed Directed Graphs

Occasionally relations are measured in which the ties can be interpreted as being either positive or negative in affect, evaluation, or meaning. For example, one might measure the relations "loves" and "hates" among the people in a group, or the relations "is allied with" and "is at war with" among countries. Such relations can be represented as a signed graph, or as a signed directed graph. We begin by defining a signed graph, and then generalize to a signed directed graph. Signed graphs and signed directed graphs are important in the study of balance and clusterability (discussed in Chapter 6).

4.4.1 Signed Graph

A *signed graph* is a graph whose lines carry the additional information of a valence: a positive or negative sign. A signed graph consists of three sets of information: a set of nodes, $\mathscr{N} = \{n_1, n_2, \ldots, n_g\}$, a set of lines, $\mathscr{L} = \{l_1, l_2, \ldots, l_L\}$, and a set of valences (or signs), $\mathscr{V} = \{v_1, v_2, \ldots, v_L\}$, attached to the lines. As usual, each line is an unordered pair of distinct nodes, $l_k = (n_i, n_j)$. But now, associated with each line is a valence, v_k either "+" or "−". A line, $l_k = (n_i, n_j)$ is assigned the valence $v_k = +$ if the tie between actors i and j is positive in meaning or affect, and a valence $v_k = -$ if the tie between the actors represented by the nodes is negative. We denote a signed graph as $\mathscr{G}_{\pm}(\mathscr{N}, \mathscr{L}, \mathscr{V})$, or simply $\mathscr{G}_{\pm}$.

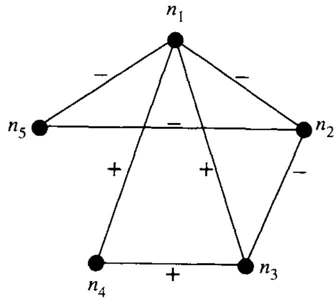
For example, we can represent alliances and hostilities among nations using a signed graph by letting nodes represent countries, and letting signed lines represent whether pairs of countries are at war with each other, "−", or have a treaty with each other, "+".

A *complete signed graph* is a signed graph in which all unordered pairs of nodes are included in the set of lines. Since all lines are present in a complete signed graph, and all lines have a valence either "+" or "−", each unordered pair of nodes is assigned either "+" or "−".

Dyads and Triads. In a signed graph, each dyad is in one of three states: There is a positive line between them, there is a negative line between them, or there is no line between them. In a complete signed graph each dyad is in one of two states, either "+" or "−".

In a complete signed graph, a triad may be in one of four possible states, depending on whether zero, one, two, or three positive (or negative) lines are present among the three nodes.

Cycles. Many properties of signed graphs (such as balance and clusterability) depend on cycles and properties of cycles. In this section we define the *sign* of a cycle in a signed graph. Recall that a cycle is a closed walk in which all nodes except the beginning and ending node are distinct. Notice that each line in a cycle in a signed graph is either "+" or "−". In a signed graph, the *sign* of a cycle is defined as the product of



$l_1 = (n_1\ n_2)$	$v_1 = -$
$l_2 = (n_1\ n_3)$	$v_2 = +$
$l_3 = (n_1\ n_4)$	$v_3 = +$
$l_4 = (n_1\ n_5)$	$v_4 = -$
$l_5 = (n_2\ n_3)$	$v_5 = -$
$l_6 = (n_2\ n_5)$	$v_6 = -$
$l_7 = (n_3\ n_4)$	$v_7 = +$
Cycle	Sign of cycle
$n_1\ n_2\ n_5\ n_1$	$- \times - \times - = -$
$n_3\ n_4\ n_1\ n_3$	$+ \times + \times + = +$
$n_1\ n_2\ n_3\ n_1$	$- \times - \times + = +$

Fig. 4.21. Example of a signed graph

the signs of the lines included in the cycle; where the sign of the product is defined as:

- $(+)(+) = +$
- $(+)(-) = -$
- $(-)(-) = +$

In brief, if a cycle has an even number of negative, "$-$", lines, then its sign is positive. However, if a cycle has an odd number of negative lines, its sign is negative.

Figure 4.21 gives an example of a signed graph and some of its cycles.

4.4.2 Signed Directed Graphs

It is straightforward to extend the idea of a signed graph to a *signed directed graph.* A signed directed graph is a directed graph in which

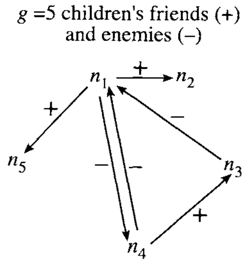


Fig. 4.22. Example of a signed directed graph

the arcs have the additional information of a positive or negative sign. A signed digraph consists of three sets of information: a set of nodes, $\mathscr{N} = \{n_1, n_2, \ldots, n_g\}$, a set of arcs, $\mathscr{L} = \{l_1, l_2, \ldots, l_L\}$, and a set of valences, $\mathscr{V} = \{v_1, v_2, \ldots, v_L\}$, attached to the arcs. In a signed directed graph, each arc is an ordered pair of distinct nodes, $< n_i, n_j >$. Associated with each arc is a valence, either "+" or "−". Since the arc $l_k =< n_i, n_j >$ is distinct from the arc $l_m =< n_i, n_j >$, the sign v_k may be different from the sign v_m. We can denote a signed directed graph as $\mathscr{G}_{d\pm}(\mathscr{N}, \mathscr{L}, \mathscr{V})$, or simply $\mathscr{G}_{d\pm}$.

Claims of friendship and enmity among people can be represented as a signed directed graph. Nominations of friends might be represented by a "+" and nominations of enemies might be represented by a "−". Figure 4.22 contains an example of a signed digraph, which we can take to represent such friendship and enmity nominations among people.

Semicycles. In a signed directed graph the most general cycles are usually referred to as *semicycles.* Recall that a semicycle is a closed sequence of distinct nodes and arcs in which each node is either adjacent to or adjacent from the previous node in the sequence. Thus a semicycle is a cycle in which the arcs may point in either direction. The *sign of a semicycle* is the product of the signs of the arcs in it.

This idea is important for studying balance and clusterability in signed directed graphs (see Chapter 6).

Signed graphs and signed directed graphs generalize graphs and directed graphs by allowing the lines or arcs to have valences. Now, let us generalize even further by allowing the lines or arcs to have other (usually numerical) values.

4.5 Valued Graphs and Valued Directed Graphs

Often social network data consist of valued relations in which the strength or intensity of each tie is recorded. Examples of valued relations include the frequency of interaction among pairs of people, the dollar amount of trade between nations, or the rating of friendship between people in a group. Such relations cannot be fully represented using a graph or a directed graph, since lines or arcs in a graph or directed graph are only present or absent (dichotomous: 0 or 1). Thus, the next step in the generalization of graphs and digraphs is to add a *value* or *magnitude* to each line or arc. Valued graphs are the appropriate graph theoretic representation for valued relations. In this section we define and describe valued graphs.

There are several special valued graphs; for example, weighted graphs and integer weighted graphs (Roberts 1976), nets and networks (Harary 1969), and Markov chains. We will briefly describe each. Concepts and definitions for valued graphs are not as well developed as they are for graphs and directed graphs; thus, our discussion of valued graphs will be briefer than our discussion of graphs and directed graphs.

A *valued graph* or a *valued directed graph* is a graph (or digraph) in which each line (or arc) carries a value. A valued graph consists of three sets of information: a set of nodes, $\mathcal{N} = \{n_1, n_2, \ldots, n_g\}$, a set of lines (or arcs), $\mathcal{L} = \{l_1, l_2, \ldots, l_L\}$, and a set of values, $\mathcal{V} = \{v_1, v_2, \ldots, v_L\}$, attached to the lines (or arcs). Associated with each line (in a graph) or each arc (in a digraph) is a value from the set of real numbers (Flament 1963). We denote a valued graph by $\mathcal{G}_V(\mathcal{N}, \mathcal{L}, \mathcal{V})$, or simply $\mathcal{G}_V$. Roberts (1976) refers to a valued digraph as a *weighted digraph*.

A valued graph represents a nondirectional valued relation, such as the number of interactions observed between each pair of people in a group. The number of interactions between actor i and actor j is the same as the number of interactions between actor j and actor i. In a valued graph the line between node n_i and node n_j is identical to the line between node n_j and node n_i ($l_k = (n_i, n_j) = (n_j, n_i)$), and thus there is only a single value, v_k, for each unordered pair of nodes.

A valued directed graph represents a directional valued relation, such as the dollar amount of manufactured goods exported from each country to each other country. Country i may export a different amount of manufactured goods to country j than country j exports to country i. In a valued directed graph, the arc from node n_i to node n_j is not the same as the arc from node n_j to node n_i ($l_k =< n_i, n_j > \neq l_m =< n_j, n_i >$), and thus there are two distinct values, one for each possible arc for the ordered pair of nodes. In general, for $l_k =< n_i, n_j >$ and $l_m =< n_j, n_i >$, v_k does not necessarily equal v_m.

Some authors allow the values to be non-numerical (for example, letters or colors). Harary, Norman, and Cartwright (1965) refer to such a valued graph as a *network.*

Special cases of valued graphs and valued directed graphs place restrictions on the possible values that the lines or arcs can take. Harary (1969) refers to a valued graph in which all values are from the *positive* real numbers as a *network* (note how a variety of authors differ in their definition of the term "network"). If all values in a valued digraph are from the set of integers, then it is what Roberts (1976) refers to as an *integer weighted digraph.*

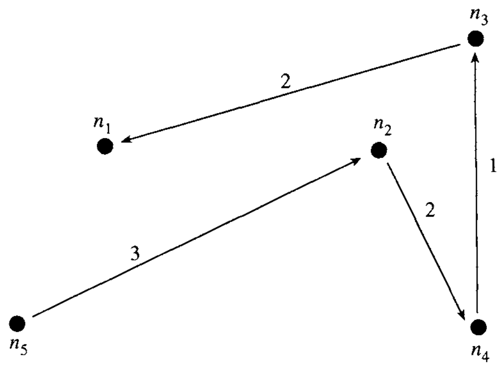
One can also consider a signed graph in which positive lines have the value $+1$ and negative lines have the value -1 as an integer-weighted graph, with integer values $+1$ and -1. A signed graph is thus a special case of a valued graph in which the values are only $+1$ and -1. Similarly, a graph is a special case of a valued graph in which each and every line has a value equal to 1.

One specific application of valued graphs that has been studied extensively is the set of graphs whose values are probabilities. These graphs are known as Markov chains, and their corresponding sociomatrices are often referred to as transition matrices or stochastic matrices (Harary 1959b). In a Markov chain the values of all arcs incident from each node are constrained to sum to 1, for all n_i, $\sum v_k = 1$ for all $l_k =< n_i, n_j >$, $j = 1, 2, \ldots, g$; further, $0 \leq v_k \leq 1$.

Often we will restrict our attention to relations that are discrete-valued, and thus can be represented as integer-weighted graphs or integer-weighted digraphs, where the values are from the non-negative integers. In this case, the value of an arc in a digraph (or a line in a graph) takes on the values $m = 1, 2, \ldots, C$.

As another example, if nominations of three best friends and three worst enemies were requested, ties might be labeled $+3$ for a best friend, $+2, +1, -1, -2$, and -3 for a worst enemy.

Figure 4.23 gives an example of a valued digraph. This figure lists the arcs and their values. For example, the arc $l_4 =< n_5, n_2 >$ has a value of 3, so $v_4 = 3$.



$l_1 = < n_3, n_1 >$ $v_1 = 2$

$l_2 = < n_2, n_4 >$ $v_2 = 2$

$l_3 = < n_4, n_3 >$ $v_3 = 1$

$l_4 = < n_5, n_2 >$ $v_4 = 3$

Fig. 4.23. Example of a valued directed graph

4.5.1 *Nodes and Dyads*

Nodes in Valued Graphs. Each node in a valued graph can have a number of lines incident with it. Similarly, each node in a valued digraph can have a number of arcs incident to it and/or from it. To each line or arc is attached a value. In a graph or digraph, nodal degree is equal to the number of lines incident with the node or the number of arcs incident to it or from it. The idea of degree does not generalize well to valued graphs, since one must consider the values attached to the lines.

One way to generalize the notion of degree to valued graphs and digraphs is to average the values over all lines incident with a node, or all arcs incident to or from a node. Such a measure reflects the average value of the lines incident with the node or of the arcs to or from the node.

Dyads in Valued Graphs. A dyad in a valued graph has a line between nodes with a specific strength. A dyad in a valued directed graph has arcs between the nodes. Each of the two arcs $< n_i, n_j >$ and $< n_j, n_i >$ has a value, which we denote by v_k and v_m. These values most likely will be different. It is of interest in such settings to compare the v_k to v_m. Models for such dyads are discussed in Chapter 15.

4.5.2 Density in a Valued Graph

In a graph or digraph, density, Δ, is defined as the ratio of the number of lines or arcs present to the maximum possible that could arise. Another way to view the density of a graph or a digraph is as the average of the values assigned to the lines/arcs. Each line or arc is given a value of 1, and pairs of nodes for which lines are absent are given a value of 0. The sum of these values is equal to the number of lines or arcs; one then divides this sum by its maximum possible value.

To generalize the notion of density to a valued graph or digraph, one can average the values attached to the lines/arcs across all lines/arcs. Thus, for a valued graph/digraph, the density is $\Delta = \sum v_k/g(g-1)$ where the sum is taken over all k. This measures the average strength of the lines/arcs in the valued graph/digraph.

4.5.3 ○Paths in Valued Graphs

Walks and paths in valued graphs are defined the same way as they are in graphs (as an alternating sequence of nodes and lines beginning and ending with nodes). However, in a valued graph (or valued digraph) since the lines (or arcs) have values attached to them, concepts such as reachability of a pair of nodes, length of a path, and distance between a pair of nodes become more complicated. In order to define these concepts for valued graphs, we must consider the values attached to each of the lines (or arcs) in a path. As Peay (1980) has noted, there are a number of different, and reasonable, ways to define distance and values for paths in a valued graph. The choice of which definition to use depends on the interpretation of the lines (arcs) and values in the graph. As in a graph, nodes n_i and n_j are *reachable* if there is a path between them. In a valued graph we can also consider "strengths" or "values" of reachability.

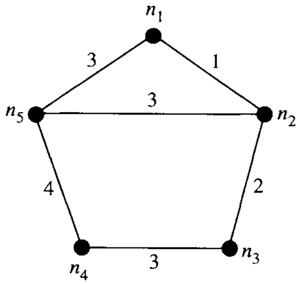
Value of a Path. The *value* of a path (semipath) is equal to the smallest value attached to any line (arc) in it (Peay 1980). Formally, the value of $W = l_1, l_2, \ldots, l_k$ from n_i to n_j equals $\min(v_1, v_2, \ldots, v_k)$. The value of a path is thus the "weakest link" in the path. This idea makes most sense if larger values indicate stronger ties. For example, if the lines represent the amount of communication between each pair of people in a group, then the value of a path between two people represents the most "restricted" amount of communication between any pair of people in the path.

Now, for simplicity, consider a valued graph in which the values attached to the lines are discrete and ordinal, and take on values $1, 2, \ldots, C$ (this is a simplifying condition that is not necessary). We define a *path at level c* as a path between a pair of nodes such that each and every line in the path has a value greater than or equal to c; that is, $v_l \geq c$ for all v_l in the path (Doreian 1969, 1974). In general, paths that include only lines with large values will have higher path values, whereas paths that include lines with small values will have lower path values. Since all values in a path at level c are greater than or equal to c, a path at level c is also a path at all values less than or equal to c. This concept is used to study cohesive subgroups for valued graphs (Chapter 7).

Reachability. We can generalize reachability for a pair of nodes to strengths of reachability in a valued graph (Doreian 1974). Consider all paths between a pair of nodes. Each of these paths has a value. The higher the value, the "stronger" the lines included in the path. In a valued graph, two nodes are *reachable at level c* if there is a path at level c between them. In other words, if two nodes are reachable at level c then there is at least one path between them that includes no line with a value less than c. If two nodes are reachable at level c, then they are reachable at any value less than c.

Path Length. If the values attached to the lines (or arcs) can be thought of as "costs" associated with the tie (such as the amount of time required to go from point i to point j), then it is useful to define the length of a path as the sum of the values of the lines in it. Flament (1963) defines the *length* of a path in a valued graph as equal to the sum of the values of the lines included in the path. If all values are equal to 1, then this definition is equivalent to the definition of path length for a graph or a directed graph since the sum is simply the number of lines (arcs) in the path. One possible problem with this quantification of path length in a valued graph is that a high value for a path can result either if the values of the lines in the path are high, or if the path is long (and thus contains many lines).

Figure 4.24 gives an example of a valued graph. It also gives the lengths and values of some paths in this graph.



$l_1 = (n_1\ n_2)$ $v_1 = 1$
$l_2 = (n_1\ n_5)$ $v_2 = 3$
$l_3 = (n_2\ n_3)$ $v_3 = 2$
$l_4 = (n_2\ n_5)$ $v_4 = 3$
$l_5 = (n_3\ n_4)$ $v_5 = 3$
$l_6 = (n_4\ n_5)$ $v_6 = 4$

Path	*Length*	*Value*
$n_1\ n_5\ n_4$	7	3
$n_1\ n_2\ n_3\ n_4$	6	1
$n_1\ n_5\ n_2\ n_3\ n_4$	11	2

Fig. 4.24. Paths in a valued graph

In the previous sections we discussed graphs (for representing dichotomous nondirectional relations) and described graphs that generalize graphs in two different ways. Directed graphs are used for representing directional relations and generalize graphs by considering the direction of the arcs between pairs of nodes. Both graphs and directed graphs represent dichotomous relations. The second way to generalize graphs (and directed graphs) is to allow the lines (or arcs) to carry values. Signed and valued graphs and digraphs generalize graphs by removing the restriction that lines (arcs) be either present or absent. A third way to generalize graphs and digraphs is to have more than one relation measured on a pair of nodes. We consider this generalization next.

4.6 Multigraphs

So far, we have discussed simple graphs, where there is at most one line between a pair of nodes. A simple graph is the appropriate representation for a social network in which a single relation is measured. When there is more than one relation, a *multigraph* is the appropriate graph theoretic representation. A multigraph, or a *multivariate (directed) graph* is a generalization of a simple graph or digraph that allows more than one set of lines (Flament 1963).

If more than one relation is measured on the same set of actors, then the graph representing this network must allow each pair of nodes to be connected in more than one way. For example, for Krackhardt's high-tech managers, each person was asked with whom they were "friends," and from whom they sought advice on the job. That is, two relations were measured on the set of actors.

A multigraph $\mathscr{G}$ consists of a set of *nodes*, $\mathscr{N} = \{n_1, n_2, \ldots, n_g\}$, and two or more sets of lines, $\mathscr{L}^+ = \{\mathscr{L}_1, \mathscr{L}_2, \ldots, \mathscr{L}_R\}$. We let R be the number of sets of lines in the multigraph, and we subscript the lines to denote to which set it belongs. If each relation is nondirectional, each line in each of the R sets is an unordered pair of distinct nodes, $l_{kr} = (n_i, n_j)$. A pair of nodes may be included in more than one set of lines. Since there are R sets of lines, each unordered pair of nodes may have from 0 up to R lines between them.

Returning to the definition of a simple graph, a graph is called *simple* if it contains no loops (or self-choices) and if each pair of nodes is joined by 0 or 1 lines. If a graph contains loops and/or any pair of nodes is adjacent via *more than one* line the graph is *complex*. Much of graph theory concentrates on simple graphs, and most of the graph theoretic concepts network researchers use pertain to simple graphs. Accordingly, most of the methods that we discuss in later chapters focus on simple graphs.

Graphs and directed graphs consider *pairs* of nodes. Lines and arcs are defined as pairs or ordered pairs of nodes. The next generalization of graphs is to consider ties among subsets of nodes.

4.7 ⊗Hypergraphs

Some social network applications consider ties among subsets of actors in a network, such as the tie among people who belong to the same club or civic organization. Such networks, called affiliation networks, or membership networks, require considering subsets of nodes in a graph, where these subsets can be of any size. Hypergraphs are the appropriate representations for such networks.

An affiliation network is a two-mode network consisting of a set of actors and a set of events. Each event is a subset of the actors from

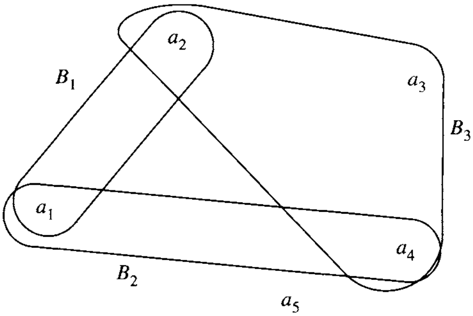


Fig. 4.25. Example of a hypergraph

$\mathscr{N}$. Thus, affiliation network data cannot be fully represented in terms of pairwise ties, since the subsets can include more than two actors. A *hypergraph*, rather than a graph, is the appropriate representation for affiliation network data.

Formally, a *hypergraph* consists of a set of objects and a collection of subsets of objects, in which each object belongs to at least one subset, and no subset is empty (Berge 1989). The objects are called *points* and the collections of objects are called *edges.* In general, for point set $\mathscr{A} = \{a_1, a_2, \ldots, a_g\}$, and edge set $\mathscr{B} = \{B_1, B_2, \ldots, B_h\}$, the hypergraph is denoted by $\mathscr{H} = (\mathscr{A}, \mathscr{B})$.

An important feature of a hypergraph is that it can also be described by the *dual hypergraph*, denoted $\mathscr{H}^\star$, by reversing the roles of the points and the edges. In general, if the hypergraph $\mathscr{H} = (\mathscr{A}, \mathscr{B})$ has point set $\mathscr{A}$ and edge set $\mathscr{B}$, then the dual hypergraph $\mathscr{H}^\star = (\mathscr{B}, \mathscr{A})$ has point set $\mathscr{B}$ and edge set $\mathscr{A}$.

To distinguish between points and edges, we introduce the following notation. When the elements in $\mathscr{A}$ are viewed as points and the elements in $\mathscr{B}$ are viewed as edges (as in hypergraph $\mathscr{H} = (\mathscr{A}, \mathscr{B})$) we will denote the elements of $\mathscr{A}$ using lowercase letters: $\{a_1, a_2, \ldots, a_g\}$, and the edges from set $\mathscr{B}$ using uppercase letters: $\{B_1, B_2, \ldots, B_h\}$. In the dual hypergraph, $\mathscr{H}^\star = (\mathscr{B}, \mathscr{A})$ with elements in $\mathscr{B}$ as points and elements in $\mathscr{A}$ as edges, we will denote the elements in $\mathscr{A}$ using uppercase letters: $\{A_1, A_2, \ldots, A_g\}$ and the elements in $\mathscr{B}$ using lowercase letters: $\{b_1, b_2, \ldots, b_h\}$.

The hypergraph in Figure 4.25 has point set $\mathscr{A} = \{a_1, a_2, a_3, a_4\}$, and edge set $\mathscr{B} = \{B_1, B_2, B_3\}$. This hypergraph could represent four actors attending three social events.

The hypothetical example in Figure 4.25 can be described in terms of each edge in $\mathscr{B}$ and the subset of points in $\mathscr{A}$ that it includes:

$$\begin{aligned} B_1 &= \{a_1, a_2\} \\ B_2 &= \{a_1, a_4\} \\ B_3 &= \{a_2, a_3, a_4\} \end{aligned}$$

Alternatively, we can list each element in $\mathscr{A}$ as an edge, and the elements in $\mathscr{B}$ as the points contained in it:

$$\begin{aligned} A_1 &= \{b_1, b_2\} \\ A_2 &= \{b_1, b_3\} \\ A_3 &= \{b_3\} \\ A_4 &= \{b_2, b_3\} \end{aligned}$$

Hypergraphs are more general than graphs. A graph is a special case of a hypergraph in which the number of points in each edge is exactly equal to two. Any graph can be represented as a hypergraph, by letting the nodes in the graph be the points in the hypergraph, and letting each line $l_k = (n_i, n_j)$ in the graph be an edge in the hypergraph. Each edge thus contains exactly two points.

4.8 Relations

Social networks are often described using formal mathematical notation as *mathematical relations* (Hage and Harary 1983; Fararo 1973; Pattison 1993; and others). We now describe this representation of a social network.

4.8.1 Definition

A mathematical relation focuses on the ordered pairs of actors in a network between whom a substantive tie is present. Relations are widely used in algebraic methods (Chapter 11).

Consider a set of objects, $\mathscr{N} = \{n_1, n_2, \ldots, n_g\}$. In a social network these objects are the actors. In a graph, these are the nodes. In a social network, ties link pairs of actors. Thus, we focus on ordered pairs of objects from $\mathscr{N}$.

The Cartesian product of two sets (or of a set with itself) is a useful mathematical entity for studying relations. The Cartesian product of two sets, $\mathscr{M}$ and $\mathscr{N}$, is the collection of all ordered pairs in which the first element in the pair belongs to set $\mathscr{M}$ and the second element belongs to set $\mathscr{N}$. We denote the Cartesian product of sets $\mathscr{M}$ and $\mathscr{N}$ as $\mathscr{M} \times \mathscr{N}$. If there are h elements in $\mathscr{M}$ and g elements in $\mathscr{N}$, and each of the h elements in $\mathscr{M}$ is paired with each of the g elements in $\mathscr{N}$, then there are $h \times g$ elements in the Cartesian product of $\mathscr{M}$ and $\mathscr{N}$.

Now, consider the Cartesian product of a set with itself: $\mathscr{N} \times \mathscr{N}$. This Cartesian product consists of all ordered pairs of objects from $\mathscr{N}$. If the set $\mathscr{N}$ is the set of all actors in a network, then the Cartesian product $\mathscr{N} \times \mathscr{N}$ is the set of all ordered pairs of actors. Thus, $\mathscr{N} \times \mathscr{N}$ is the collection of all ordered pairs of actors for whom substantive ties could be present. For example, if we are studying friendship among people, and the set of actors is $\mathscr{N}$, then the Cartesian product, $\mathscr{N} \times \mathscr{N}$, is the set of all possible ordered pairs of people, in which the first person could choose the second person as a friend. However, friendship ties are usually present between only some of the ordered pairs of people.

A *relation*, R, on the set $\mathscr{N}$ is defined as a subset of the Cartesian product $\mathscr{N} \times \mathscr{N}$ (Hage and Harary 1983). In substantive terms, the relation R consists of all ordered pairs $< n_i, n_j >$ for whom the substantive tie from i to j is present. Relations are conveniently expressed using algebraic notation (see Chapter 3). If the ordered pair $< n_i, n_j > \in R$ then we write iRj.

4.8.2 Properties of Relations

There are several important properties of relations: reflexivity, symmetry, and transitivity. Unlike a simple graph that excludes loops, a relation allows the possible $< n_i, n_i >$ tie to be present. A relation is *reflexive* if all possible $< n_i, n_i >$ ties are present in R; that is, iRi for all i. If *no* $< n_i, n_i >$ ties are present in R, then the relation is *irreflexive*. If a relation is neither reflexive nor irreflexive, then it is *not reflexive* (Fararo 1973; and Hage and Harary 1983). A relation that is not reflexive is one on which iRi for some but not all i.

Symmetry is another property of a relation. A relation is *symmetric* if it has the property that iRj if and only if jRi, for all i and j. That is, the ordered pair $< n_i, n_j > \in R$ if and only if $< n_j, n_i > \in R$. A symmetric relation is one in which all dyads are either mutual or null.

On some relations the presence of the $< n_i, n_j >$ tie implies the absence of the $< n_j, n_i >$ tie. Such a relation is *antisymmetric*. An antisymmetric relation is one on which iRj implies that not jRi. An example of an antisymmetric relation is the relation "beats" in a sporting tournament if each team plays each other team no more than once. If team i beats team j, then it cannot be the case that team j beats team i (if they play only once and no ties are allowed).

A relation that is neither symmetric nor antisymmetric is called *not symmetric*, *non-symmetric*, or *asymmetric*. A relation that is not symmetric is one for which iRj and jRi, for some but not all i and j.

A third important property of a relation is whether or not it is transitive. A relation is *transitive* if whenever iRj and jRk, then iRk, for all i, j, and k. Substantively, transitivity captures the notion that "a friend of a friend is a friend." If person i "chooses" person j as a friend, and person j in turn "chooses" person k as a friend, then, if friendship is transitive, person i will "choose" person k as a friend.

We will return to these important properties below and show how matrices can be used to study symmetry, reflexivity, and transitivity in social networks.

4.9 Matrices

The information in a graph $\mathscr{G}$ may also be expressed in a variety of ways in matrix form. There are two such matrices that are especially useful. The first is the *sociomatrix*, first introduced in Chapter 3. The second is the *incidence* matrix. We will begin by describing these matrices for a single nondirectional relation (or graph), and then generalize to matrices for directional relations (digraphs), valued relations (valued graphs), and hypergraphs.

4.9.1 Matrices for Graphs

The Sociomatrix. The primary matrix used in social network analysis is called the *adjacency* matrix, or *sociomatrix*, and is denoted by **X**. Graph theorists refer to this matrix as an *adjacency matrix* because the entries in the matrix indicate whether two nodes are adjacent or not. In the study of social networks, the adjacency matrix is usually referred to as a sociomatrix. Sociomatrix is the term we will use most often.

A sociomatrix is of size $g \times g$ (g rows and g columns) for one-mode networks. There is a row and column for each node, and the rows and columns are labeled $1, 2, \ldots, g$. The rows and columns index nodes in the graph, or actors in the network, in identical order. The entries in the sociomatrix, x_{ij}, record which pairs of nodes are adjacent. In the

Table 4.2. *Example of a sociomatrix: "lives near" relation for six children*

X

	n_1	n_2	n_3	n_4	n_5	n_6
n_1	-	0	0	0	1	1
n_2	0	-	1	0	0	0
n_3	0	1	-	0	0	0
n_4	0	0	0	-	1	1
n_5	1	0	0	1	-	1
n_6	1	0	0	1	1	-

sociomatrix, there is a 1 in the (i, j)th cell (row i, column j) if there is a line between n_i and n_j, and a 0 in the cell otherwise. In other words, if nodes n_i and n_j are adjacent, then $x_{ij} = 1$, and if nodes n_i and n_j are not adjacent, then $x_{ij} = 0$.

For the present, we are focusing on graphs where the lines are not directed and are neither signed nor valued. That is, a line between two nodes is either present or it is absent. If a line is present, it goes both from n_i to n_j and from n_j to n_i; thus, $x_{ij} = 1$, and $x_{ji} = 1$.

The sociomatrix for a graph (for a nondirectional relation) is *symmetric*. A matrix is symmetric if $x_{ij} = x_{ji}$ for all i and j; thus the entries in the upper right and lower left triangles are identical. The entries on the diagonal, values of x_{ii}, are udefined, if we do not allow loops in the graph.

The sociomatrix for a complete graph contains 1's in all off-diagonal cells. Since all nodes are adjacent, $x_{ij} = x_{ji} = 1$ for all $i \neq j$. The sociomatrix for an empty graph contains 0's everywhere, since no nodes are adjacent.

For example in Figure 4.1, nodes n_2 and n_3 are adjacent, since the line $l_1 = (n_2, n_3)$ is in the set of lines $\mathscr{L}$. Thus, $x_{23} = 1$ and $x_{32} = 1$ in the sociomatrix. Nodes n_1 and n_3 are not adjacent, since there is no line between the two, therefore $x_{13} = 0$ and $x_{31} = 0$.

The sociomatrix for the graph in Figure 4.1 is given in Table 4.2. Note that the diagonal entries are undefined, since we are focusing on simple graphs, those without loops; that is, x_{ii} is undefined if there are no loops. Also note the entries are binary, since a line is either present, $x_{ij} = 1$, or absent, $x_{ij} = 0$, between any two ñodes. Thus, the sociomatrix for a graph contains only 1's and 0's, since pairs of nodes are either adjacent, or not. Finally, note the matrix is symmetric, since a line between n_i and n_j is also a line between n_j and n_i, so $x_{ij} = x_{ji}$.

Table 4.3. *Example of an incidence matrix: "lives near" relation for six children*

I	l_1	l_2	l_3	l_4	l_5	l_6
n_1	1	1	0	0	0	0
n_2	0	0	1	0	0	0
n_3	0	0	1	0	0	0
n_4	0	0	0	1	1	0
n_5	1	0	0	1	0	1
n_6	0	1	0	0	1	1

In summary, the sociomatrix records for each pair of nodes whether the nodes are adjacent or not. The next matrix we describe records which nodes are incident with which lines.

The Incidence Matrix. The second matrix that can be used to present the information in a graph is called the *incidence matrix*, **I**, or $\mathbf{I}(\mathscr{G})$, and records which lines are incident with which nodes. The incidence matrix has nodes indexing the rows, and lines indexing the columns. Since there are g nodes and L lines, the matrix **I** is of size $g \times L$; there is a row for every node, and a column for every line. The matrix entry I_{ij} equals 1 if node n_i is incident with line l_j, and equals 0 if node n_i is not incident with line l_j. Since the line $l_k = (n_i, n_j)$ is incident with the two nodes n_i and n_j, each column in **I** has exactly two 1's in it, recording the two nodes incident with the line.

The incidence matrix is binary, since a line is either incident with a node or it is not. However, it is not necessarily square.

The incidence matrix for the graph in Figure 4.1 is given in Table 4.3. Note that since there are $g = 6$ nodes and $L = 6$ lines, **I** is 6×6.

The sociomatrix and the incidence matrix both contain all the information in a graph. The set of nodes and lines in a graph is completely described by the information in either matrix.

4.9.2 Matrices for Digraphs

The sociomatrix, **X**, of a digraph has elements x_{ij} equal to 1 if there is an arc from row node n_i to column node n_j, and 0 otherwise. The value in cell x_{ij} is equal to 1 if the arc $< n_i, n_j >$ is in $\mathscr{L}$. That is, the entry in the (i, j)th cell of **X** is equal to 1 if the actor represented by row node n_i "chooses" the actor represented by column node n_j. Since the "choice"

Table 4.4. *Example of a sociomatrix for a directed graph: friendship at the beginning of the year for six children*

X

	n_1	n_2	n_3	n_4	n_5	n_6
n_1	-	1	0	0	1	0
n_2	0	-	1	0	0	1
n_3	0	1	-	0	0	0
n_4	0	0	0	-	1	0
n_5	0	0	0	0	-	1
n_6	0	1	0	0	0	-

from i to j is substantively different from the "choice" from j to i, the entry in x_{ij} may be different from the entry in x_{ji}. For example, if actor i "chose" actor j, but j did not reciprocate, there would be a 1 in the x_{ij} cell, and a 0 in the x_{ji} cell.

The sociomatrix for the digraph in Figure 4.16 (the relation is friendship at the beginning of the school year) is given in Table 4.4. Note that, for example, the mutual choices between actors Drew (n_2) and Sarah (n_6) are represented by a 1 in both the x_{26} and x_{62} cells of this sociomatrix.

4.9.3 Matrices for Valued Graphs

A valued graph can also be represented as a sociomatrix. The entry in cell x_{ij} is the value associated with the line between node n_i and node n_j in a valued graph, or the value associated with the arc from n_i to n_j in a valued directed graph.

The sociomatrix for a valued graph (representing a valued nondirectional relation) has entries, x_{ij}, that record the value v_k associated with the line or arc l_k between n_i and n_j. For an undirected valued graph, there is a single value, v_k, associated with the line $l_k = (n_i, n_j)$, and thus the value in cell (i, j) is equal to the value in cell (j, i); $x_{ij} = x_{ji} = v_{ij}$. However, for a directed valued graph the arc $l_k =< n_i, n_j >$ with value v_k and the arc $l_m =< n_j, n_i >$ with value v_m are distinct. Thus, $x_{ij} = v_k$ and $x_{ji} = v_m$, which may differ. The entry in cell (i, j) of **X** records the strength of the tie from actor i to actor j.

4.9.4 Matrices for Two-Mode Networks

For two-mode networks the sociomatrix is of size $g \times h$, where the rows label the nodes in $\mathcal{N} = \{n_1, n_2, \ldots, n_g\}$ and the columns label the nodes in $\mathcal{M} = \{m_1, m_2, \ldots, m_h\}$.

4.9.5 ○Matrices for Hypergraphs

The matrix for a hypergraph, denoted by **A**, is a g by h matrix that records which points are contained within which edges. For the hypergraph, $\mathcal{H}(\mathcal{N}, \mathcal{M})$, with point set $\mathcal{N} = \{n_1, n_2, \ldots, n_g\}$ and edge set $\mathcal{M} = \{M_1, M_2, \ldots, M_h\}$, the matrix $\mathbf{A} = \{a_{ij}\}$ has an entry $a_{ij} = 1$ if point n_i is in edge M_j, and 0 otherwise. The matrix **A** has been called the *incidence* matrix for the hypergraph (Berge 1989), since it codes which points are incident with which edges.

The sociomatrix is the most common form for presenting social network data. It is especially useful for computer analyses. In addition, it is a very flexible representation since graphs, directed graphs, signed graphs and digraphs, and valued graphs and digraphs can all be represented as sociomatrices.

4.9.6 Basic Matrix Operations

In this section we describe and illustrate basic matrix operations that are used in social network analysis.

Vocabulary. The *size* of a matrix (also called its *order*) is defined as the number of rows and columns in the matrix. A matrix with g rows and h columns is of size g by h, or equivalently $g \times h$. A sociomatrix for a network with a single set of actors and one relation has g rows and g columns, and is thus of size $g \times g$. If a matrix has the same number of rows and columns, it is *square*. Otherwise, it is *rectangular*. A sociomatrix for a single set of actors and a single relation is necessarily square.

Each entry in a matrix is called a *cell*, and is denoted by its row index and column index. So, cell x_{ij} is in row i and column j of the matrix.

For a square matrix, the *main diagonal* of the matrix consists of the entries for which the index of the row is equal to the index of the column, that is, $i = j$. Thus, the main diagonal contains the entries in the x_{ii} cells, for $i = 1, 2, \ldots, g$. In a sociomatrix, the entries on the main diagonal are the self-"choices" of actors in the network, or the loops in the graph. If these are undefined, as they are when we exclude loops from a graph or do not measure self-choices of actors in the network, then the entries on the main diagonal of a sociomatrix are undefined. In this instance, we will put a "–" in the (i, i)th diagonal entry of a sociomatrix.

An important property of a square matrix is whether it is *symmetric*. A matrix is symmetric if $x_{ij} = x_{ji}$, for all cells. If this is not true, then the matrix is not symmetric, that is, if there are some cells where $x_{ij} \neq x_{ji}$. The sociomatrix for a graph (representing a nondirectional relation) is symmetric, since the line (n_i, n_j) is identical to the line (n_j, n_i), and thus $x_{ij} = x_{ji}$ for all i and j. However, the sociomatrix for a digraph (representing a directional relation) is not necessarily symmetric, since the arc $< n_i, n_j >$ is not the same as the arc $< n_i, n_j >$, and thus the entry in cell x_{ij} is not necessarily the same as the entry in cell x_{ji}.

We now turn to some important matrix operations, including matrix permutation, the transpose of a matrix, matrix addition and subtraction, matrix multiplication, and Boolean matrix multiplication.

Matrix Permutation. In a graph the assignment of numbers to the nodes is arbitrary. The only information in the graph is which pairs of nodes are adjacent. Similarly, in a sociomatrix, the order of the rows and columns indexing the actors in the network or the nodes in the graph is arbitrary, so long as the rows and columns are indexed in the same order. Any rearrangement of rows, and simultaneously of columns, of the sociomatrix does not change the information about adjacency of nodes, or ties between actors. Sometimes it is useful to rearrange the rows and columns in the sociomatrix to highlight patterns in the network. For example, if the relation represented in a sociomatrix is advice-seeking among managers in several departments in a corporation, then it might be useful to place managers in the same department next to each other in the rows and columns of the sociomatrix in order to study advice-seeking within departments.

A *permutation* of a set of objects is any reordering of the objects. If a set contains g objects, then there are $g! = g \times (g-1) \times (g-2) \times \cdots \times 1$ possible permutations of these objects. For example, there are $3 \times 2 \times 1 = 6$ permutations of the integers $\{1, 2, 3\}$. Thus, there are six ways to rearrange the rows and columns of a sociomatrix for three actors, simply by relabeling (simultaneously) the rows and columns.

Matrix permutations can be used in the study of cohesive subgroups (Chapter 7), and are especially important in constructing blockmodels (Chapter 10), and in evaluating the goodness-of-fit of blockmodels (Chapter 16). Matrix permutations are also useful if the graph is bipar-

Table 4.5. *Example of matrix permutation*

X

	n_1	n_2	n_3	n_4	n_5
n_1	-	0	1	0	1
n_2	0	-	0	1	0
n_3	1	0	-	0	1
n_4	0	1	0	-	0
n_5	1	0	1	0	-

X permuted

	n_5	n_1	n_3	n_2	n_4
n_5	-	1	1	0	0
n_1	1	-	1	0	0
n_3	1	1	-	0	0
n_2	0	0	0	-	1
n_4	0	0	0	1	-

tite. Recall that the nodes in a bipartite graph can be partitioned so that all lines are between nodes in different subsets. Thus, it is helpful to permute the rows and columns of the sociomatrix so that nodes in the same subset are in rows (and columns) that are next to each other in the sociomatrix.

Sometimes the patterns of ties between actors is not clear until we permute both the rows and the columns of the matrix. For example in Table 4.5, an arbitrary labeling of nodes might have ordered the rows (and columns) n_1, n_2, n_3, n_4, n_5, as in the sociomatrix at the top of the table. However, the permutation at the bottom of the table has the nodes in the order: $5, 1, 3, 2, 4$, there are now 1's in the upper left and lower right corners of the sociomatrix. With this new order of rows and columns, it is clear that ties are present among the nodes represented by rows and columns $5, 1$, and 3 and among nodes represented by rows and columns 2 and 4, but there are no ties between these two subsets. This pattern of two separate subsets was difficult, if not impossible, to see in the original sociomatrix.

Transpose. The *transpose* of a matrix is constructed by interchanging the rows and columns of the original matrix. For matrix **X** we denote its transpose as $\mathbf{X}'$ with entries $\{x'_{ij}\}$. For matrix **X**, the elements of its transpose $\mathbf{X}'$ are $x'_{ij} = x_{ji}$.

If a matrix, **X**, is symmetric, then **X** and its transpose, $\mathbf{X}'$, are identical; $\mathbf{X} = \mathbf{X}'$. Thus, the matrix for a graph (representing a nondirectional

Table 4.6. *Transpose of a sociomatrix for a directed relation: friendship at the beginning of the year for six children*

$\mathbf{X}'$

	n_1	n_2	n_3	n_4	n_5	n_6
n_1	-	0	0	0	0	0
n_2	1	-	1	0	0	1
n_3	0	1	-	0	0	0
n_4	0	0	0	-	0	0
n_5	1	0	0	1	-	0
n_6	0	1	0	0	1	-

relation) is always identical to its transpose, since $x_{ij} = x_{ji}$ for all i and j. However, the matrix for a digraph (representing a directional relation) is not necessarily identical to its transpose, since the sociomatrix for a directional relation is not, in general, symmetric.

The transpose of a sociomatrix is substantively interesting since it is analogous to reversing the direction of the ties between pairs of actors. In a sociomatrix, an entry of 1 in cell (i, j) indicates that there is a tie from row actor i to column actor j. In the transpose of the sociomatrix, a 1 in cell (i, j) indicates that row actor i *received* a tie from column actor j. For a directional relation represented as a directed graph, the transpose of the sociomatrix represents the converse of the directed graph; $x'_{ij} = 1$ if $x_{ji} = 1$.

Table 4.6 gives the transpose of the sociomatrix in Table 4.4.

Addition and Subtraction. The addition of two matrices of the same size (the same number of rows and columns) is defined as the sum of the elements in the corresponding cells of the matrices. For matrices $\mathbf{X}$ and $\mathbf{Y}$, both of size g by h, we define $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$, where $z_{ij} = x_{ij} + y_{ij}$.

Similarly we can define matrix subtraction as the difference between the elements in the corresponding cells of the matrices. For matrices $\mathbf{X}$ and $\mathbf{Y}$, both of size g by h, we define $\mathbf{Z} = \mathbf{X} - \mathbf{Y}$, where $z_{ij} = x_{ij} - y_{ij}$.

Matrix Multiplication. Matrix multiplication is a very important operation in social network analysis. It can be used to study walks and reachability in a graph, and is the basis for compounding relations in the analysis of relational algebras (see Chapter 11).

Consider two matrices: $\mathbf{Y}$ of size $g \times h$, and $\mathbf{W}$ of size $h \times k$. The number of columns in $\mathbf{Y}$ must equal the number of rows in $\mathbf{W}$. We define

YW = Z

$$\mathbf{Y} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 3 & 2 \end{bmatrix} \quad \mathbf{W} = \begin{bmatrix} 0 & 2 \\ 1 & 1 \\ 2 & 3 \end{bmatrix} \quad \mathbf{Z} = \begin{bmatrix} 2 & 5 \\ 7 & 11 \end{bmatrix}$$

$$z_{11} = (1 \times 0) + (0 \times 1) + (1 \times 2) = 0 + 0 + 2 = 2$$
$$z_{12} = (1 \times 2) + (0 \times 1) + (1 \times 3) = 2 + 0 + 3 = 5$$
$$z_{21} = (1 \times 0) + (3 \times 1) + (2 \times 2) = 0 + 3 + 4 = 7$$
$$z_{22} = (1 \times 2) + (3 \times 1) + (2 \times 3) = 2 + 3 + 6 = 11$$

Fig. 4.26. Example of matrix multiplication

the product of two matrices as $\mathbf{Z} = \mathbf{YW}$ where the elements of $\mathbf{Z} = \{z_{ij}\}$ are equal to:

$$z_{ij} = \sum_{l=1}^{h} y_{il} w_{lj}. \tag{4.11}$$

The matrix product **Z** has g rows and k columns. The value in cell (i, j) of **Z** is equal to the sum of the products of corresponding elements in the ith row of **Y** and the jth column of **W**.

Figure 4.26 gives an example of matrix multiplication. The first matrix in the product, **Y**, is of size 2×3, and the second matrix, **W**, is of size 3×2. Hence, the product, **Z**, is of size 2×2.

Powers of a Matrix. Now, consider the sociomatrix **X** of size g by g. We denote the product of a matrix times itself, **XX** as $\mathbf{X}^2$, with entries $x_{ij}^{[2]}$. Since there are g rows and g columns in **X** there are also g rows and g columns in $\mathbf{X}^2$.

Multiplying $\mathbf{X}^2$ by the original sociomatrix, **X**, yields the matrix $\mathbf{X}^3 = \mathbf{XXX}$. In general, we define $\mathbf{X}^p$ (**X** to the pth power) as the matrix product of **X** times itself, p times.

Table 4.7 shows a matrix and some of its powers.

Boolean Matrix Multiplication. The result of multiplying two matrices, say **X** and **Y**, is a new matrix, **Z**, with entries whose values are defined by equation (4.11). In many social network applications it is sufficient to consider only whether these entries are non-zero. Such arithmetic is usually referred to as *Boolean.* Boolean matrix multiplication yields the Boolean product of two matrices, which we denote by $\mathbf{Z}^{\otimes} =$

$\mathbf{X} \otimes \mathbf{Y}$. The entries of a Boolean product are either 0 or 1, and are defined as:

$$z_{ij}^{\otimes} = \begin{cases} 1 & \text{if} \quad \sum_{l=1}^{h} y_{il} w_{lj} > 0 \\ 0 & \text{if} \quad \sum_{l=1}^{h} y_{il} w_{lj} = 0. \end{cases}$$

Thus Boolean matrix multiplication results in values that are equal to 1 if regular matrix multiplication results in a non-zero entry, and equal to 0 otherwise. Boolean multiplication is the basis for constructing relational algebras (Chapter 11), and can be used to study walks and reachability in graphs.

4.9.7 Computing Simple Network Properties

Now, let us see how these matrix operations can be used to study some graph theoretic concepts. We will first describe how to use matrix multiplication to study walks and reachability in a graph and then show how properties of matrices can be used to quantify nodal degree and graph density.

Walks and Reachability. Matrix operations can be used to study walks and reachability in both graphs and directed graphs.

Graphs. first, let us consider the sociomatrix for a graph (representing a nondirectional relation). As defined in equation (4.11), the value $x_{ij}^{[2]} = \sum_{k=1}^{g} x_{ik} x_{kj}$. The product $x_{ik} x_{kj}$, one term in this sum, is equal to 1 only if both $x_{ik} = 1$ and $x_{kj} = 1$. In terms of the graph, $x_{ik} x_{kj} = 1$ only if both lines (n_i, n_k) and (n_k, n_j) are present in $\mathscr{L}$. If this is true, then the walk $n_i n_k n_j$ is present in the graph. Thus, the sum $\sum_{k=1}^{g} x_{ik} x_{kj}$ counts the number of walks of length 2 between nodes n_i and n_j, for all k. The entries of $\mathbf{X}^2 = \{x_{ij}^{[2]}\}$ give exactly the number of walks of length 2 between n_i and n_j.

Similarly, we can consider walks of any length by studying powers of the matrix $\mathbf{X}$. For example, elements of $\mathbf{X}^3$ count the number of walks of length 3 between each pair of nodes. Such multiplications can be used to find walks of longer lengths. In general, the entries of the matrix $\mathbf{X}^p$ (the matrix $\mathbf{X}$ raised to the pth power) give the total number of walks of length p from node n_i to node n_j.

Recall that two nodes are reachable if there is a path (and thus, a walk) between them. Since every path is a walk, we can study reachability of pairs of nodes by considering the powers of the matrix $\mathbf{X}$ that count walks of a given length. Also, recall that the longest possible path in a graph is equal to $g - 1$ (any path longer than $g - 1$ must include some node(s) more than once, and so is not a path). Thus, if two nodes are reachable, then there is at least one path (and thus at least one walk) of length $g - 1$ or less between them.

Consider now whether there is a walk of length k or less between two nodes, n_i and n_j. If there is a walk of length k or less, then, for some value of $p \leq k$, the element $x_{ij}^{[p]}$ will be greater than or equal to 1. One way to determine whether two nodes are reachable is to examine all matrices, $\{\mathbf{X}^p, 1 \leq p \leq g-1\}$. If two nodes are reachable, then there is a non-zero entry in one or more of the matrices of this set. When these product matrices are summed, for $p = 1, 2, \ldots, (g-1)$, we obtain a matrix,

$$\mathbf{X}^{[\Sigma]} = \mathbf{X} + \mathbf{X}^2 + \mathbf{X}^3 + \ldots + \mathbf{X}^{g-1}$$

whose entries give the total number of walks from n_i to n_j, of any length less than or equal to $g - 1$. Since any two nodes that are reachable are connected by a path (and thus a walk) of length $g - 1$ or less, non-zero entries in the matrix $\mathbf{X}^{[\Sigma]}$ indicate pairs of nodes that are reachable. A 0 in cell (i, j) of $\mathbf{X}^{[\Sigma]}$ means that there is no walk between nodes n_i and n_j, and thus these two nodes are not reachable.

It is useful to define a *reachability* matrix, $\mathbf{X}^{[R]} = \{x_{ij}^{[R]}\}$, that simply codes for each pair of nodes whether they are reachable, or not. The entry in cell (i, j) of $\mathbf{X}^{[R]}$ is equal to 1 if nodes n_i and n_j are reachable, and equal to 0 otherwise. We can calculate these values by looking at the elements of $\mathbf{X}^{[\Sigma]}$, and noting which ones are non-zero. Non-zero elements of $\mathbf{X}^{[\Sigma]}$ indicate reachability; hence, we define

$$x_{ij}^{[R]} = \begin{cases} 1 & \text{if } x_{ij}^{[\Sigma]} \geq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{4.12}$$

The elements of $\mathbf{X}^{[R]}$ indicate whether nodes n_i and n_j are reachable or not.

Directed Graphs. Now, consider matrix products of sociomatrices for directed graphs. These products will allow us to study directed walks and reachability for directed graphs.

First, look at the entries in $\mathbf{X}^2$. If $\mathbf{X}$ is a sociomatrix for a directed graph, then $x_{ij} = 1$ means that the arc $< n_i, n_j >$ is in $\mathscr{L}$. As usual, the value of the product $x_{ik}x_{kj}$ is equal to 1 if both $x_{ik} = 1$ and $x_{kj} = 1$. In the directed graph, $x_{ik}x_{kj} = 1$ only if both arcs $< n_i, n_k >$ and $< n_k, n_j >$

are present in $\mathscr{L}$. If this is true, then the directed walk $n_i \rightarrow n_k \rightarrow n_j$ is present in the graph. The sum $\sum_{k=1}^{g} x_{ik}x_{kj}$ thus counts the number of directed walks of length 2 beginning at node n_i and ending at node n_j, for all k. Thus, the entries of $\mathbf{X}^2 = \{x_{ij}^{[2]}\}$ give exactly the number of directed walks of length 2 from n_i to n_j.

Similarly, we can consider directed walks of any length by studying powers of the matrix $\mathbf{X}$. In general, the entries of the matrix $\mathbf{X}^p$ (the pth power of the sociomatrix for a directed graph) give the total number of directed walks of length p beginning at row node n_i and ending at column node n_j.

As with the powers of the sociomatrix for a graph, when the product matrices, $\mathbf{X}^p$, are summed, for $p = 1, 2, \ldots, (g-1)$, we obtain a matrix, denoted by $\mathbf{X}^{[\Sigma]}$, whose entries give the total number of directed walks from row node n_i to column node n_j, of any length less than or equal to $g - 1$.

We can also define the reachability matrix for a directed graph, $\mathbf{X}^{[R]} = \{x_{ij}^{[R]}\}$, that codes for each pair of nodes whether they are reachable, or not. The entry in cell (i, j) of $\mathbf{X}^{[R]}$ is equal to 1 if there is a directed path from row node n_i to column node n_j, and equal to 0 otherwise. If $x_{ij}^{[R]} = 1$ then node n_j is reachable from node n_i. Since directed paths consist of arcs all "pointing" in the same direction, there may be a directed path from node n_i to node n_j (thus $x_{ij}^{[R]} = 1$), without there necessarily being a directed path from node n_j to node n_i (thus $x_{ji}^{[R]}$ could $= 0$). Thus, the reachability matrix for a directed graph is not, in general, symmetric.

Geodesics and Distance. The (geodesic) distance from n_i to n_j can be found by inspecting the power matrices. The distance from one node to another is the length of a shortest path between them. In a graph, this distance is the same from n_i to n_j as it is from n_j to n_i. In a digraph, these distances can be different.

These distances are sometimes arrayed in a *distance matrix*, with elements $d(i, j)$. To find these distances using matrices, focus on the (i, j) elements of the power matrices, starting with $p = 1$. When $p = 1$, the power matrix is the sociomatrix, so that if $x_{ij} = 1$, the nodes are adjacent, and the distance between the nodes equals 1. If $x_{ij} = 0$ and $x_{ij}^{[2]} > 0$, then there is a shortest path of length 2. And so forth. Consequently, the first power p for which the (i, j) element is non-zero gives the length of the shortest path and is equal to $d(i, j)$. Mathematically, $d(i, j) = min_p x_{ij}^{[p]} > 0$.

Table 4.7. *Powers of a sociomatrix for a directed graph*

X

	n_1	n_2	n_3	n_4	n_5	n_6
n_1	-	1	0	0	1	0
n_2	0	-	1	0	0	1
n_3	0	1	-	0	0	0
n_4	0	0	0	-	1	0
n_5	0	0	0	0	-	1
n_6	0	1	0	0	0	-

$\mathbf{X}^2$

	n_1	n_2	n_3	n_4	n_5	n_6
n_1	0	0	1	0	0	2
n_2	0	2	0	0	0	0
n_3	0	0	1	0	0	1
n_4	0	0	0	0	0	1
n_5	0	1	0	0	0	0
n_6	0	0	1	0	0	1

$\mathbf{X}^3$

	n_1	n_2	n_3	n_4	n_5	n_6
n_1	0	3	0	0	0	0
n_2	0	0	2	0	0	2
n_3	0	2	0	0	0	0
n_4	0	1	0	0	0	0
n_5	0	0	1	0	0	1
n_6	0	2	0	0	0	0

$\mathbf{X}^4$

	n_1	n_2	n_3	n_4	n_5	n_6
n_1	0	0	3	0	0	3
n_2	0	4	0	0	0	0
n_3	0	0	2	0	0	2
n_4	0	0	1	0	0	1
n_5	0	2	0	0	0	0
n_6	0	0	2	0	0	2

$\mathbf{X}^5$

	n_1	n_2	n_3	n_4	n_5	n_6
n_1	0	6	0	0	0	0
n_2	0	0	4	0	0	4
n_3	0	4	0	0	0	0
n_4	0	2	0	0	0	0
n_5	0	0	2	0	0	2
n_6	0	4	0	0	0	0

The diameter of a graph or digraph is the length of the largest geodesic in the graph or digraph. If the graph is connected or if the digraph is at least strongly connected, the diameter of the graph is then the largest entry in the distance matrix; otherwise, the diameter is infinite or undefined.

Computing Nodal Degrees. In this section we describe how to calculate nodal degree from the matrices associated with graphs and directed graphs. We first describe calculations of nodal degree for a graph, and then nodal indegree and outdegree for a directed graph.

Graphs. Recall that the degree of a node, $d(n_i)$, is equal to the number of lines incident with the node in the graph. Nodal degrees may be found by summing appropriate elements in either the sociomatrix or in the incidence matrix. In the incidence matrix $\mathbf{I}$, with elements $\{I_{ij}\}$, the degrees of the nodes are equal to the row sums, since the rows correspond to nodes and the entries are 1 for every line incident with the row node. Specifically,

$$d(n_i) = \sum_{j=1}^{L} I_{ij}.$$

Each row contains as many 1's as there are lines incident with the node in that row. Thus, summing over columns (that is, lines) gives the number of lines incident with the node.

In the sociomatrix $\mathbf{X}$ for a graph (representing a nondirectional relation) the nodal degrees are equal to either the row sums or the column sums. The ith row or column total gives the degree of node n_i:

$$d(n_i) = \sum_{j=1}^{g} x_{ij} = \sum_{i=1}^{g} x_{ij} = x_{i+} = x_{+j}. \tag{4.13}$$

Directed Graphs. Now consider the indegrees and outdegrees of nodes in a directed graph. Recall that the indegree of a node is the number of nodes incident *to* the node (the number of arcs terminating at it) and the outdegree of a node is the number of arcs incident *from* the node (the number of arcs originating from it). Notice that row i of a sociomatrix contains entries $x_{ij} = 1$ if node n_j is incident from node i. The number of 1's in row i is thus the number of nodes incident *from* node n_i, and is equal to the *outdegree* of node n_i. Similarly, the entries in column i of a sociomatrix contain entries $x_{ji} = 1$ if node n_j is incident *to* node n_i. Thus, the number of 1's in column i is equal to the *indegree* of node n_i. The row totals of $\mathbf{X}$ are equal to the nodal outdegrees, and the column totals of $\mathbf{X}$ are equal to the nodal indegrees. Formally,

$$d_O(n_i) = \sum_{j=1}^{g} x_{ij} = x_{i+}, \tag{4.14}$$

and

$$d_I(n_i) = \sum_{j=1}^{g} x_{ji} = x_{+i}. \tag{4.15}$$

Computing Density. The density of a graph, digraph, or valued (di)graph can be calculated as the sum of all entries in the matrix, divided by the possible number of entries:

$$\Delta = \frac{\sum_{i=1}^{g}\sum_{j=1}^{g} x_{ij}}{g(g-1)}. \tag{4.16}$$

4.9.8 Summary

We have showed how many of the graph theoretic properties for nodes, pairs of nodes, and the graph as a whole can be calculated using matrix representations. These representations are quite useful, as Katz (1947) first realized.

4.10 Properties of Graphs, Relations, and Matrices

In this chapter we have noted three important properties of social networks: reflexivity, symmetry, and transitivity. In this section, we show how they can be studied by examining matrices, relations, and graphs.

4.10.1 Reflexivity

In our discussion of graphs we have focused on simple graphs, which, by definition, exclude loops. Thus, a simple graph is irreflexive, since no $< n_i, n_i >$ are present. On occasion, however, one may wish to allow loops. In that case, if all loops are present, the graph represents a reflexive relation. In a sociomatrix loops are coded by the entries along the main diagonal of the matrix, x_{ii} for all i. A relation is reflexive if, in the sociomatrix, $x_{ii} = 1$ for all i. An irreflexive relation has entries on the main diagonal of the sociomatrix that are undefined. Finally, a relation that is not reflexive (also not irreflexive) has some, but not all, values of $x_{ii} = 1$. In terms of a directed graph, some, but not all, $< n_i, n_i >$ arcs are present.

4.10.2 Symmetry

A relation is symmetric if, whenever i "chooses" j, then j also "chooses" i; thus, iRj if and only if jRi. A nondirectional relation (represented by a graph) is always symmetric. In a directed graph symmetry implies that whenever the arc $l_k =< n_i, n_j >$ is in the set of lines $\mathscr{L}$, the arc $l_m =< n_j, n_i >$ is also in $\mathscr{L}$. In other words, dyads are either null or mutual. The sociomatrix for a symmetric relation is symmetric; $x_{ij} = x_{ji}$ for all distinct i and j. If the matrix $\mathbf{X}$ is symmetric, then it is identical to its transpose, $\mathbf{X}'$; $x_{ij} = x'_{ij}$ for all i and j.

4.10.3 Transitivity

Transitivity is a property that considers patterns of triples of actors in a network or triples of nodes in a graph. A relation is transitive if every time that iRj and jRk, then iRk. If the relation is "is a friend of," then the relation is transitive if whenever i "chooses" j as a friend and j "chooses" k as a friend, then i "chooses" k as a friend.

Transitivity can be studied by considering powers of a sociomatrix. Recall that $\mathbf{X}^{[2]} = \mathbf{XX}$ codes the number of walks of length 2 between each pair of nodes in a graph; thus, an entry $x_{ij}^{[2]} \geq 1$ if there is a walk $n_i \to n_k \to n_j$ for at least one node n_k. Thus, in order for the relation to be transitive, whenever $x_{ij}^{[2]} \geq 1$, then x_{ij} must equal 1.

One can check for transitivity of a relation by comparing the square of a sociomatrix with the sociomatrix. Thus, a transitive relation is noteworthy in that ties present in $\mathbf{X}^2$ are a subset of the ties present in $\mathbf{X}$.

4.11 Summary

Graph theory is a useful way to represent network data. Actors in a network are represented as nodes of a graph. Nondirectional ties between actors are represented as lines between the nodes of a graph. Directed ties between actors are represented as arcs between the nodes in a digraph. The valences of ties are represented by a "+" or "−" sign in a signed graph or digraph. The strength associated with each line or arc in a valued graph or digraph is assigned a value. Many of the concepts of graph theory have been used as the foundation of many theoretical concepts in social network analysis.

There are many, many references on graph theory. We recommend the following texts. Harary (1969) and Bondy and Murty (1976) are excellent mathematical introductions to graph theory, with coverage ranging from proofs of many of the statements we have made, to solutions to a variety of applied problems. The excellent text by Frank (1971) is more mathematically advanced and focuses on social networks. Similarly, the classic text by Harary, Norman, and Cartwright (1965) is also focused on directed graphs, and is quite accessible to beginners. Roberts (1976, 1978) and Hage and Harary (1983) provide very readable, elementary introductions to graph theory, with many concepts illustrated on anthropological network data. In their introduction to network analysis, Knoke and Kuklinski (1982) also describe some elementary ideas in graph theory. Ford and Fulkerson (1962), Lawler (1976), Tutte (1971), and others give mathematical treatments of special, advanced topics in graph theory, such as theories of matroids and optimization of network configurations. The topic of tournaments is treated in the context of paired comparisons by David (1988). A more mathematical discussion of tournaments can be found in Moon (1968). Berge (1989) discusses hypergraphs in detail.