1 · Digraphs and Structures

It is only the very great and good who have any living faith in the simplest axioms.

SAMUEL BUTLER, Erewhon

The word "structure" is found extensively in the literature of the social sciences. "Social structure" and such related concepts as "kinship structure," "authority structure," "communication structure," and "sociometric structure" are commonplace.\(^1\) Psychologists speak of such matters as "personality structure," "cognitive structure," and "attitude structure."\(^2\) Linguists are interested in the "structure of a language" or "syntactical structure."\(^3\) And many other examples could be cited. But despite the widespread use of structural concepts in the social sciences, it is fair to say that the formal analysis of structure has been relatively underdeveloped in these fields. The technical terminology employed in describing structures is meager; few concepts are defined rigorously. As a consequence, the social scientific description of structural properties tends to be couched in ambiguous terminology, and detailed studies of structure, as such, are rather rare.

In the natural sciences and engineering, problems of "structure" have also been encountered. As a result of these and other less applied considerations, mathematicians have actively turned their attention to certain branches of mathematics which deal with the abstract notion of structure. One line of investigation began with the work of Euler

\(^1\) See, for example, Bavelas (1950), Glanzer and Glaser (1959, 1961), Hunter (1953), Lindeley and Borgatta (1954), Merion (1957), Nadel (1957), and Weiss (1956).

\(^2\) For examples, see Baldwin (1942), Lewin (1951), Rosenberg et al. (1960), and Sanford (1956).

\(^3\) See Harris (1951) and Chomsky (1957).
(1707–1783) in the fields subsequently known as topology and graph theory. It has been significantly advanced by such men as Cayley (who was interested in certain structural problems in chemistry), Kirchhoff (whose laws of electrical network theory are famous), and more recently by numerous engineers concerned with communication systems. It has also benefited from the contributions of mathematicians and logicians who have worked on problems in the theory of relations, matrices, and combinatorics. The resulting body of mathematical knowledge contains, we believe, much that will be of value to social scientists in their investigations of various kinds of empirical structures.

Our object in this volume, then, is to present some mathematical theory about the abstract notion "structure." Specifically, we shall provide an introduction to the theory of directed graphs, or more briefly "digraphs" (a term suggested by G. Pólya). This theory is concerned with patterns of relationships among pairs of abstract elements. As such, digraph theory makes no reference to the empirical world. Nevertheless, it has potential usefulness to the empirical scientist, for it can serve as a mathematical model of the structural properties of any empirical system consisting of relationships among pairs of elements.

Consider, for example, the "communication structure" of a group of people. We may conceive of each member of the group as an "element" and the fact that a particular member can communicate directly to another as a "relationship." Then, upon coordinating these empirical entities and relationships to the abstract terms of digraph theory, we obtain a digraph which represents the communication structure of the group. The properties of this digraph, with which digraph theory is concerned, are at the same time properties of the communication structure.

In a similar way, a "sociometric structure" may be conceived as consisting of elements, which are people, and relationships among pairs of people, as indicated by interpersonal choices obtained from a questionnaire. A digraph may be constructed to represent any particular sociometric structure.

As another example, consider an individual's "preference structure" on a set of objects as revealed by the method of paired comparisons. In this case, the elements of the structure are objects, for example, pictures, and the relationships are the individual's indicated preferences of each object relative to each other. Again, the resulting structure may be represented by a digraph.

It should be apparent that there is an almost endless variety of empirical structures which may be represented by digraphs. The structure of a task may be conceived in terms of required precedence
relationships on a set of operations. The authority structure of an organization may be thought of as supervisory relationships on a set of positions. Or, the causal structure of a scientific theory may be viewed as causal relationships on a set of variables.

Knowledge of digraph theory is useful to the researcher interested in the structural properties of any empirical system, for it provides concepts, theorems, and methods appropriate to the analysis of structure, *per se*. There are three principal benefits which the scientist may gain from employing digraph theory in his treatment of structural phenomena.

First, he will find that his vocabulary for describing empirical structures is enriched by useful new terms having precise meanings, for the language of digraphs contains a large number of concepts which refer to relatively complex structural properties. We shall find in the course of this volume that precise definitions can be given to such ideas as the degree of connectedness of a structure, its diameter, its vulnerability, and its stratification into levels. Although everyday language contains terms referring to such properties, these are poorly defined and have no clear conceptual relationships to one another. Even though this use of digraph theory exploits its definitions more than its theorems, the contribution is a significant one.

Second, digraph theory and associated branches of mathematics provide techniques of computation and formulas for calculating certain quantitative features of empirical structures. Matrix algebra has an especially close relation to digraph theory, and we shall examine in some detail its computational value in dealing with structures. We shall see, for example, that it is possible to ascertain the distance in a structure from one element to another and to construct measures of such things as the degree of centrality and the relative status of an element within a structure. Other indexes will be developed to measure the degree of consistency and the degree of balance of a structure as a whole. Digraph theory thus provides the researcher with useful means for quantifying certain aspects of structure.

The third, and perhaps most important, benefit of the use of digraph theory stems from its theorems. We shall find that the axioms for digraph theory to be presented in this chapter lead to an extensive body of logically derived statements. Each of these statements or theorems becomes a valid assertion about any empirical structure that satisfies the axioms of digraph theory. The theorems thereby give additional information about empirical structures. By specifying properties of digraphs that necessarily follow from given conditions, they permit us to draw conclusions about certain properties of a structure from knowledge about other properties. Since it is difficult to illustrate the place
of theorems in digraph theory before its nature has been elaborated, we defer further discussion of this topic until the end of the chapter.

In order for the researcher to gain these benefits from using digraph theory, it is essential that he have unambiguous rules for coordinating basic terms of digraph theory to empirical phenomena so that its axioms, and hence its derived theorems, produce, when thus interpreted, true statements about the empirical world. Thus, we may hope to benefit from the precision of digraph theory, but this gain requires skill in handling empirical data. In particular, we must have adequate operational procedures for identifying relevant empirical entities and relationships among them. If these are unequivocally identified, then digraph theory specifies structural properties that must be found in the empirical world.

With this general orientation, let us turn now to a consideration of the basic nature of digraph theory.

NETS AND RELATIONS

It is best to begin with fundamentals. The theory of digraphs is based on an axiom system consisting of four primitives (undefined terms), together with four axioms (or postulates) which give us an understanding of the primitives and of their relations to one another. As we shall see, digraph theory can be thought of as growing out of more general mathematical theories about nets and relations. All three theories have the same primitives, but the theory of nets uses only two of the axioms whereas the theory of relations uses three. We shall begin with nets, proceed to relations, and then consider digraphs.

Before stating the primitives and axioms for a net, we describe them more intuitively. The diagram shown in Figure 1.1 is a very simple net \( N \), consisting of two points and a directed line. In the picture of a net, the points are conventionally depicted by dots labeled \( v_1, v_2, \ldots, v_r \) and the directed lines by arcs labeled \( x_1, x_2, \ldots, x_r \). The direction of each line is indicated by an arrowhead. In Figure 1.1, the direction of the line \( x_1 \) can also be described by saying that its first point, denoted \( f(x_1) \), is \( v_1 \) and its second point, \( s(x_1) \), is \( v_2 \).
The four primitives of nets (and also of relations and of digraphs) are:

- **P₁**: A set \( V \) of elements called “points.”
- **P₂**: A set \( X \) of elements called “directed lines,” or more briefly, “lines.”
- **P₃**: A function \( f \) whose domain is \( X \) and whose range is contained in \( V \).
- **P₄**: A function \( s \) whose domain is \( X \) and whose range is contained in \( V \).

The first two of these primitives are self-explanatory. The second two relate the lines to the points by means of two functions \( f \) and \( s \) which serve to identify the “first” and the “second” point of each line, respectively. In the net of Figure 1.1, \( V = \{ v₁, v₂ \} \) and \( X = \{ x₁ \} \). The image \( f(x₁) \) of the function \( f \) is \( v₁ \), the first point of \( x₁ \). The image \( s(x₁) \) of the function \( s \) is \( v₂ \), the second point of \( x₁ \). It is in this sense that the line \( x₁ \) is “directed” from \( v₁ \) to \( v₂ \). In general, for any line \( x \) of \( X \), the image \( fx \) of the function \( f \) is called the first point of \( x \) and the image \( sx \) of the function \( s \) is the second point of \( x \). The line \( x \) is incident with both its points \( fx \) and \( sx \). Thus every line of a net is directed from its first point to its second point.

The axioms for a net are:

- **A₁**: The set \( V \) is finite and not empty.
- **A₂**: The set \( X \) is finite.

The first axiom excludes consideration of an empty net with no points at all and of a net with an infinite number of points. Then the second axiom avoids nets with a finite number of points but an infinite number of lines. Thus these two axioms impose no restrictions on the structure of a net other than the number of its points and lines.

Figure 1.2(a) shows a net \( N \) which is a bit more complicated than the previous one. In this case, \( N \) has three points and six lines. We see that the first point of line \( x₁ \) is \( v₁ \) and its second point is \( v₂ \); that is, \( fx₁ = v₁ \) and \( sx₁ = v₂ \). We also find that \( fx₃ = fx₁ = v₁ \) and \( sx₃ = sx₁ = v₂ \). Moreover, \( fx₅ = sx₃ = v₂ \) and \( fx₆ = sx₆ = v₂ \).

Although this picture is drawn on a plane and resembles figures found in Euclidean geometry, a net is quite a different object. For example, the lengths of the arcs and the location of the dots on the page have no significance in the theory of nets. Thus, the net shown in Figure 1.2(b) is the same as that in Figure 1.2(a) even though it looks rather different. All of the statements in the preceding paragraph about part (a) apply equally to part (b) of this figure. In thinking about nets (and digraphs) we shall have to break certain habits built up over years of experience with Euclidean geometry.
It will be useful now to make certain observations about the lines which a net may contain. A line $x$ of a net is called a loop if $fx = xx$, in other words, if it has the same first and second point. The net of Figure 1.2 has two loops: $x_3$ and $x_6$. We say that two lines $x_i$ and $x_j$ are parallel if $fx_i = fx_j$ and $xx_i = xx_j$. Thus in the net under consideration, the lines $x_2$ and $x_3$ are parallel, as also are the loops $x_4$ and $x_5$. Note, however, that $x_3$ and $x_4$ are not parallel. When a net has parallel lines distinct meanings may be given to the different lines from one point to another. Thus, for example, the two lines from $v_1$ to $v_2$ might indicate that person $v_1$ both likes $v_2$ and tends to communicate messages to him. It is also possible to let each line indicate a unit of strength of a relationship, in which case we would say that the relationship from $v_1$ to $v_2$, perhaps that of liking, is stronger than that from $v_3$ to $v_1$ since there are two lines from $v_1$ to $v_2$ and only one from $v_3$ to $v_1$. We shall see that these complications do not arise in digraphs and shall delay further consideration of them until Chapter 14.

We may now define a relation. A relation is a net in which no two distinct lines are parallel. Clearly, every relation is a net, by definition, but not every net is a relation. Figure 1.2(a) shows a net that is not a relation, whereas Figure 1.3 shows one that is.

* In view of the axiom system, we are dealing here, of course, with a finite relation.
Since there are no parallel lines in a relation, any line $x$ is determined by specifying its first point $fx$ and its second point $sx$. And since a relation, by definition, has a finite number of points, it must necessarily have a finite number of lines as well. If $u$ and $v$ are two points of a relation and $fx = u$ and $sx = v$, we may denote the line $x$ by $(u, v)$ to indicate that $x$ is the ordered pair whose first element is $u$ and whose second element is $v$. When $(u, v)$ is in a relation $R$, we also write $uRv$. For example, in the relation $R$ of Figure 1.3, we have $e_2Rf_1$, $e_2Rv_1$, $e_2Re_2$, and $e_2Rv_2$ to stand for lines $x_1$, $x_2$, $x_3$, and $x_4$, respectively.

Properties of Relations

In this volume we shall need frequently to refer to certain properties that a relation may have. Since these are presented in most textbooks in logic, we shall treat them quite briefly here.

A relation $R$ is reflexive if every point of $R$ is on a loop. We say that $R$ is symmetric if whenever $uRv$, then $vRu$. A relation $R$ is transitive if for any three distinct points $u$, $v$, $w$ of $R$, whenever $uRv$ and $vRw$, then $uRw$. And a relation $R$ is said to be complete if for every pair of distinct points $u$ and $v$ in $R$, at least one of the ordered pairs $(u, v)$ or $(v, u)$ is in $R$.

Each of these properties has a certain kind of opposite property. We shall be concerned primarily with only two of these. A relation $R$ is irreflexive if no point of $R$ has a loop. And $R$ is said to be asymmetric if $uRv$ precludes $vRu$ for distinct points $u$ and $v$.

Any particular relation $R$ may possess various combinations of these properties. The reader should verify that the relation $R$ drawn in Figure 1.3 has none of these six properties. Consider the relation "greater than" on the set of integers. Clearly, this relation is irreflexive—

---

5 For a discussion of properties of relations, see Copi (1954, Ch. 3).
6 To be strictly consistent, we should speak here, and later, of a finite set of integers. We hope that this slight indiscernation will be forgiven.
no integer is greater than itself. It is asymmetric—for any two integers \( u \) and \( v \), if \( u \) is greater than \( v \), then \( v \) is not greater than \( u \). It is transitive—if it is given that \( u \) is greater than \( v \) and \( v \) is greater than \( w \), it follows that \( u \) is greater than \( w \). And it is complete—for any two distinct integers, one is greater than the other.

Certain combinations of properties give rise to a variety of “orders” on a set of points. A relation is called an (irreflexive) complete order if it is irreflexive, asymmetric, transitive, and complete. As we have just seen, the relation “greater than” on a set of integers has these properties and is therefore such an order.

Upon removing the requirement of completeness from this definition, we obtain a partial order, which is an irreflexive, asymmetric, and transitive relation. Consider the employees of a firm and the relation “can give orders to.” This relation is a partial order if it satisfies the following three plausible statements: (a) No one can give orders to himself; (b) if person \( u \) can give orders to \( v \), then \( v \) cannot give orders to \( u \); and (c) if \( u \) can give orders to \( v \) and \( v \) can give orders to \( w \), then \( u \) can give orders to \( w \). Since this relation is not necessarily complete, it may happen that there are two people, perhaps foremen in different departments, neither one of whom can give orders to the other.

By a reflexive complete order we mean a relation that is reflexive, asymmetric, transitive, and complete. It can be verified that the relation “greater than or equal to” on the set of integers is such an order. And a relation whose only requirement is that it be reflexive and transitive is called a quasi-order. We shall see in Chapter 2 that “reachability” is a quasi-order.

In concluding this section, we consider an especially important combination of properties. If a relation \( R \) is reflexive, symmetric, and transitive, it is known as an equivalence relation. Consider again the set of
integers discussed above. Let $R$ be the relation such that $u R v$ if and only if the difference $u - v$ is even. We see that $R$ is reflexive, since every difference $u - u = 0$, which is even. It is symmetric, for if $u - v$ is even, then $v - u$ is even. And $R$ is transitive, for if $u - v$ and $v - w$ are both even, then $u - w$ is even. By definition, then, $R$ is an equivalence relation. We observe that $R$ is not complete because no odd number is in relation $R$ to any even number. Clearly, the relation shown in Figure 1.4 is an equivalence relation. The best known equivalence relation is equality itself.

**DIGRAPHS**

We are now prepared to consider the axiom system for digraphs. Although the primitives or undefined terms of this system, and some of the axioms, are the same as those for nets and relations, we repeat them here for convenience of reference.

A digraph satisfies the following axiom system.

The primitives are:

- $P_1$: A set $V$ of elements called "points."
- $P_2$: A set $X$ of elements called "lines."
- $P_3$: A function $f$ whose domain is $X$ and whose range is contained in $V$.
- $P_4$: A function $s$ whose domain is $X$ and whose range is contained in $V$.

The axioms are:

- $A_1$: The set $V$ is finite and not empty.
- $A_2$: The set $X$ is finite.
- $A_3$: No two distinct lines are parallel.
- $A_4$: There are no loops.

Comparing the axioms for nets and for digraphs, we see that a digraph is a net with no loops and no parallel lines. In other words, a digraph is an irreflexive relation.

In discussing digraphs we shall denote points and lines in the manner developed for nets. Sometimes we refer to points by the letters $u, v, w$, and at other times by the notation $v_1, v_2, \ldots, v_n$. The lines of a digraph are often indicated by $x_1, x_2, \ldots, x_n$. When we wish to denote a line in terms of its two points, we write $uv$ for a line from $u$ to $v$ or $v_1v_2$ for a line from $v_1$ to $v_2$. In Figure 1.5 we show a digraph $D$ with its customary notation. Here the number of points in the set $V$ is $p = 3$, and the number of lines in the set $X$ is $q = 3$. The three points are

---

1 It is easily seen that axioms $A_1$ and $A_2$ together imply axiom $A_3$. Therefore $A_4$ could have been omitted from this list and derived as the first theorem.
denoted $v_1$, $v_2$, $v_3$ and the three lines are $x_1 = v_3v_1$, $x_2 = v_1v_2$, and $x_3 = v_2v_1$.

Let us summarize what we have learned thus far about nets, relations, and digraphs. Every digraph is a relation, and every relation is a net, but there are nets that are not relations and there are relations that are not digraphs. In this book we shall concentrate on digraphs. This means that we exclude from consideration loops and parallel lines. In doing so we simplify the mathematical theory and focus upon structural properties. At the end of the book we shall return to a consideration of loops and parallel lines. The exclusion of loops does not prevent our using digraphs to analyze relations that have loops. In dealing with such a relation we form its digraph simply by ignoring the loops. Nearly all the important structural properties are the same whether loops are included or not. And in many relations, such as "can communicate to," or "likes to work with," the question of whether loops are included or not makes no essential difference. Thus, there is ordinarily no difficulty in ignoring the distinction between a digraph and a relation.

By an interpretation of a digraph, we mean the coordination of empirical elements to its points and lines. In other words, an interpretation tells what entities are represented by the points and what empirical relationships by the lines. Thus, one interpretation of the digraph of Figure 1.6 might be the following: the points $v_1$, $v_2$, $v_3$ represent three individuals and each line, say $x_4 = v_4v_1$, indicates that individual $v_4$ is a neighbor of $v_1$. Since there are no lines in this digraph joining points $v_2$ and $v_3$. 
we may conclude that individuals $v_2$ and $v_3$ are not neighbors. If one were to insist that everyone is his own neighbor, then we would be dealing with a relation that is reflexive, and loops would have to be added at every point. It is clear, however, that for most purposes nothing significant is lost by treating the relation as a digraph.

Kinds of Digraphs

Since every digraph is a relation, digraphs may be described in terms of the properties of relations discussed above. The axiom system for digraphs demands only that the relation be irreflexive. If, however, certain additional properties are required, we obtain a particular kind of digraph. We now briefly describe some of the major kinds of digraphs.

A symmetric digraph is an irreflexive symmetric relation. Thus, for every line $uv$ in a symmetric digraph $D$, there is also a line $vu$. Such a digraph is shown in Figure 1.6. Symmetric digraphs constitute an important special class that correspond to graphs. Systematic treatments of the theory of graphs have been provided by König (1936), Berge (1958), and Ore (1962).

A complete symmetric digraph is one whose relation is both complete and symmetric. Therefore, every pair of points is joined by two lines, one in each direction. We denote a complete symmetric digraph by $K_n$, where $p$ is the number of points in the digraph. The digraph of Figure 1.7(a) is $K_4$, while that in Figure 1.6 is symmetric but not complete.

![Figure 1.7](image-url)

A complete asymmetric digraph is both complete and asymmetric, as for example in Figure 1.8. One interpretation of such a digraph arises if we let the points correspond to the players in a chess tournament and each line $uv$ indicate that player $u$ defeats player $v$. Then, if every player plays every other and no game ends in a draw, the digraph
representing the outcome of the tournament is complete and asymmetric. For this reason, such digraphs are commonly called tournaments, and we shall consider these structures in Chapter 11. In Figure 1.8 we see that player \( v_1 \) won all his games, whereas \( v_4 \) lost all of his.

\[ \text{FIGURE 1.8} \]

There is nothing in the axiom system for digraphs that precludes the possibility that a digraph may have no lines at all. A digraph without any lines is called \textit{totally disconnected}, as shown in Figure 1.7(b).

A \textit{transitive digraph} \( D \) is one which contains a line \( uv \) whenever lines \( uv \) and \( vw \) are in \( D \), for any distinct points \( u, v, w \). Clearly, a digraph that is transitive and asymmetric is a partial order. If, in addition, the digraph is complete, then it is a complete order. The digraph shown in Figure 1.8, having all these properties, is a complete order. It is possible, of course, for a digraph to be asymmetric and complete but not transitive. Figure 1.9 shows such a digraph representing the outcome of a tiddlywinks tournament among Harvard, Yale, and Princeton. We find that Harvard beat Yale, Yale beat Princeton, and Princeton beat Harvard. The following quotation from a review of the literature on dominance relations among birds and mammals reveals that such inconsistency is not confined to college teams. "Not infrequently, flocks contain triangular pecking relations, that is \( A \) will dominate \( B \), \( B \) will dominate \( C \), which in turn dominates \( A \)" (Collias, 1951, p. 390).

Certain empirical structures of interest to social scientists may generate a particular kind of digraph. For example, DeSoto and Kuethe (1958, 1959) have investigated the properties which people attribute to various kinds of interpersonal relations. They found, among other things, that

\[ \text{FIGURE 1.9} \]
most people believe that the following relations are likely to be symmetric: trusts, confides in, likes, dislikes, and hates. In other words, people expect structures based on these relations to correspond to symmetric digraphs. DeSoto and Kuehne also report a tendency to attribute transitivity and asymmetry to such relations as dominates and fears. Structures of this kind correspond to a partial order.

Implication Digraphs

There is an interesting interpretation of digraphs which is helpful in proving certain kinds of theorems. Consider a collection of four propositions, $p_1, p_2, p_3, p_4$. We assume that each of these is either true or false. The propositions themselves might come from any universe of discourse, say Euclidean geometry. For example,

$p_1$: $ABC$ is an equilateral triangle.
$p_2$: $ABC$ is an equiangular triangle.
$p_3$: $ABC$ is a triangle in which at least two sides are equal (an isosceles triangle).
$p_4$: $ABC$ is a triangle in which at least two angles are equal.

It is known from Euclidean geometry that propositions $p_3$ and $p_4$ are logically equivalent, that is, each implies the other. It is also known that every equilateral triangle is equiangular, and conversely. This information can be depicted by a digraph in which points represent propositions and lines indicate the relationship of implication, as shown in Figure 1.10(a).

It is immediately obvious that, in addition, $p_1$ implies $p_2$ and $p_2$ implies $p_4$. With this further information, we may construct the digraph shown in Figure 1.10(b). But it is known from propositional logic that the relation of implication among propositions is transitive. This
applies to the digraph of Figure 1.10(b), in which the lines $p_1p_3$ and $p_3p_4$ occur. Thus, it follows that $p_1$ implies $p_4$. Similarly, from the presence of lines $p_2p_4$ and $p_4p_7$, we conclude that $p_2$ implies $p_7$. All of the implications that hold for these four propositions are shown in the digraph of Figure 1.10(c).

A digraph in which the points are interpreted as propositions and the lines as implication is called an implication digraph. Therefore, each of the digraphs of Figure 1.10 is an implication digraph. In a total implication digraph there is a line from proposition $p_i$ to $p_j$ if and only if $p_i$ implies $p_j$. Thus, a total implication digraph displays all possible implications between the propositions which constitute its points. For example, Figure 1.10(c) is a total implication digraph since it is known that $p_3$ and $p_4$ do not imply $p_1$ or $p_2$. In proving theorems about digraphs, we shall often make use of (implication) digraphs themselves.

Implication digraphs are especially useful in proving that a set of propositions are all logically equivalent. If they are, we know that their total implication digraph is symmetric, transitive, and complete, since the digraph of an equivalence relation has these properties. Suppose that we have four propositions and wish to prove that they are all equivalent. It is sufficient, then, to prove just four implications—for example, those shown in Figure 1.11. Once these are established, we may use the fact that implication is a transitive relation and add a line $uw$ to the digraph whenever lines $uw$ and $uv$ have been established. Repeated applications of this procedure results in a total implication digraph that is complete, symmetric, and hence an equivalence relation.

**FIGURE 1.11**

**DIGRAPHS AND MATRICES**

In this section we show how digraphs can be represented by means of matrices. An $r \times s$ matrix is a rectangular array of $rs$ numbers called the entries of the matrix, arranged in $r$ rows and $s$ columns. We denote the entry in the $i$th row and $j$th column of a matrix $M$ by $m_{ij}$. The number $m_{ij}$ is also called the entry of the $i,j$ cell of $M$. A matrix $M$ is
often written with square brackets in the form

\[ M = \begin{bmatrix}
  m_{11} & m_{12} & \ldots & m_{1n} \\
  m_{21} & m_{22} & \ldots & m_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  m_{n1} & m_{n2} & \ldots & m_{nn}
\end{bmatrix} \]

For brevity we write: \( M = [m_{ij}] \).

Given a digraph \( D \), its adjacency matrix, \( A(D) = [a_{ij}] \), is a square matrix with one row and one column for each point of \( D \), in which the entry \( a_{ij} = 1 \) if line \( v_iv_j \) is in \( D \), while \( a_{ij} = 0 \) if \( v_i \rightarrow v_j \) is not in \( D \).

Consider a digraph of five points, \( V = \{v_1, v_2, v_3, v_4, v_5\} \), whose relation consists of the ordered pairs \((v_1, v_2), (v_1, v_3), (v_2, v_4), (v_3, v_2), \) and \((v_4, v_3)\). Figure 1.12 shows this digraph and its adjacency matrix.

\[ \begin{array}{ccccc}
  v_1 & v_2 & v_3 & v_4 & v_5 \\
  0 & 1 & 1 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 \\
\end{array} \]

\[ A(D) = \begin{bmatrix}
  0 & 1 & 1 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 \\
\end{bmatrix} \]

Row sum: 3, 0, 1, 1, 0
Column sum: 0, 2, 2, 1, 0

**FIGURE 1.12**

By an *ordering* of the points of \( D \), we mean their designation as first, second, third, and so on. It should be noted that the adjacency matrix is determined by the particular ordering of the points. Thus, if we were to take a different ordering of the points of the digraph of Figure 1.12, we might obtain a different adjacency matrix. The digraph of Figure 1.13 is the same digraph as in Figure 1.12 but the ordering of the points is
different. It can be seen that the matrix of Figure 1.13 is not the same as that of Figure 1.12. Despite the fact that a digraph may have more than one adjacency matrix, we shall usually refer to the adjacency matrix of a digraph when its points have been ordered by means of their subscripts. Therefore, when we speak of the adjacency matrix of a digraph we assume that the order of the points is understood.

\[ A(D) = \begin{bmatrix}
  v_1 & v_2 & v_3 & v_4 & v_5 \\
  v_1 & 0 & 0 & 0 & 0 \\
  v_2 & 1 & 0 & 1 & 1 \\
  v_3 & 1 & 0 & 0 & 0 \\
  v_4 & 0 & 0 & 1 & 0 \\
  v_5 & 0 & 0 & 0 & 0
\end{bmatrix} \]

Row sum
0 3 1 1 0

Column sum 2 0 2 1 0

**FIGURE 1.13**

Certain features of a digraph may be seen in its adjacency matrix. Thus, the symmetry or asymmetry of a digraph is reflected in a definite way. If a digraph is asymmetric, then the existence of a line \( v_i v_j \), precludes the existence of the line \( v_j v_i \). Therefore, it follows that in the adjacency matrix of an asymmetric digraph if \( a_{ij} = 1 \), then \( a_{ji} = 0 \). It can readily be seen that the digraph of Figure 1.13 is asymmetric. In the adjacency matrix of a symmetric digraph an entry of 1 in the \( i, j \) cell implies an entry of 1 in the \( j, i \) cell.

The row and column sums of the adjacency matrix indicate the number of lines originating and terminating at each point of the digraph. The outdegree of point \( v \), written \( \text{od}(v) \), is the number of lines from \( v \). It is evident that each row sum of the adjacency matrix gives the out-degree of the corresponding point. Thus, in the digraph of Figure 1.13 \( \text{od}(v_2) = 3 \). The indegree of point \( v \), written \( \text{id}(v) \), is the number of lines to \( v \). Clearly, each column sum of the adjacency matrix indicates the
indegree of the corresponding point.\footnote{Unlike most human beings, a point has an id but no superego.} In our example, id(v_2) = 0, there being no lines which terminate at v_2.

A point u is adjacent to v if the line uv is in D; u is adjacent from v if the line vu is in D. Thus, we may say that id(v) is the number of points adjacent to v, and that od(e) is the number of points adjacent from v. The outbundle of v is the set of lines from v; the inbundle of v is the set of lines to v. Clearly, od(e) indicates the number of lines in the outbundle of v, and id(v) indicates the number in the inbundle of v. The bundle of v is the union of the inbundle and the outbundle. These are, of course, mutually exclusive sets of lines incident with v. The total degree of v, id(v) + od(e), is the number of lines incident with v. We immediately obtain the following equation

\[ \text{id}(v) = \text{id}(v) + \text{od}(e). \]

The sum of all entries of the adjacency matrix \( A(D) \) is the sum of all the row sums of \( A \) or the sum of all the column sums of \( A \). This information is contained in the following formulas, in which \( p \) and \( q \) are the numbers of points and lines of \( D \).

**Theorem 1.1.** The sum of the indegrees of all the points of any digraph is equal to the sum of the outdegrees, and their common value is the number of lines; symbolically,

\[ \sum_{i=1}^{p} \text{id}(v_i) = q, \quad \text{and} \quad \sum_{i=1}^{q} \text{od}(e_i) = q. \]

These two equations can be expressed alternatively by stating the following formula for the average indegree, \( \overline{id} \), and the average outdegree, \( \overline{od} \), of the points of a digraph.

**Corollary 1.1a.** In any digraph, \( \overline{id} = \overline{od} = \frac{q}{p} \).

It will be useful to classify every point of a digraph according to the combination of its indegree and outdegree. An isolate is a point whose outdegree and indegree are both 0. A transmitter is a point whose outdegree is positive and whose indegree is 0. A receiver is a point whose outdegree is 0 and whose indegree is positive. A carrier is a point whose outdegree and indegree are both 1. Any other point is an ordinary point.

Each kind of point is illustrated by the digraph of Figure 1.13 and its adjacency matrix. By reference to these, the class of each point can
be determined readily:

\[ \begin{align*}
& v_3 \text{ is an isolate, } \text{od}(v_3) = \text{id}(v_3) = 0 \\
& v_2 \text{ is a transmitter, } \text{od}(v_2) = 3, \text{id}(v_2) = 0 \\
& v_1 \text{ is a receiver, } \text{od}(v_1) = 0, \text{id}(v_1) = 2 \\
& v_4 \text{ is a carrier, } \text{od}(v_4) = \text{id}(v_4) = 1 \\
& v_3 \text{ is ordinary, } \text{od}(v_3) = 1, \text{id}(v_3) = 2.
\end{align*} \]

The product of the indegree and the outdegree of a point contains information concerning its classification. For, if id(e)·od(e) = 0, then at least one of the two values, id(e) and od(e), must be 0: if they are both 0, then e is an isolate; if only id(e) = 0, then e is a transmitter; if only od(e) = 0, then e is a receiver. If the product id(e)·od(e) is not 0, then e must be in one of the two remaining classes: if this product is 1, then both id(e) and od(e) are 1, and e is a carrier; if id(e)·od(e) > 1, then both the indegree and outdegree are positive and at least one of them is greater than 1, and e is an ordinary point.

The empirical meaning of the various terms just considered may be illustrated by taking a digraph D to represent a communication network.

The outdegree of point v indicates the number of people that person v can communicate to directly, and the inbundle of v corresponds to the communication links going directly from v to other people. The indegree of point v indicates the number of people who can communicate directly to person v, and the inbundle of v corresponds to the links going directly to v from other people. With this interpretation of a digraph, an isolate point corresponds to a person who can neither send nor receive messages in the network, a transmitter to one who can send but not receive messages, a receiver to one who can receive but not send, and the two remaining types to persons who can both send and receive (a carrier being more constrained than an ordinary point).

**ISOMORPHISM OF DIGRAPHS**

It is possible, of course, for two different communication networks to have the same "structure." Suppose, for example, that there are two groups each consisting of ten people. Suppose, further, that in each group everyone can communicate directly to everyone else. The digraph corresponding to each of these communication networks consists of ten points and is complete and symmetric. These two groups are certainly different since they are composed of different individuals. Nonetheless, the communication networks of these two different groups have the same structure. Or, to give another example, suppose that both of these groups are organized in a hierarchical fashion so that the "top boss"
has three subordinates, each of whom has two subordinates. Again, the authority structures of these two different groups are the same; each has one transmitter, three ordinary points, and six receivers.

These examples are special cases of the concept of "isomorphism" of two digraphs. Two digraphs $D$ and $E$ are isomorphic if there exists a one-to-one correspondence between their points which preserves their directed lines. That is, $D$ and $E$ are isomorphic if they have the same number $p$ of points and if one can order their points respectively $v_1, v_2, \ldots, v_p$ and $u_1, u_2, \ldots, u_p$ so that for any $i$ and $j$, line $v_i v_j$ is in $D$ if and only if line $u_i u_j$ is in $E$. Such a correspondence is called an isomorphism between $D$ and $E$.

It often happens that when drawn, two isomorphic digraphs $D$ and $E$ have different appearances. In Figure 1.14 are two such digraphs. It is easy to verify that these two digraphs are isomorphic, and that corresponding points have the same subscripts. For example, we note that $v_1 v_2$ is in $D$ and that $u_1 u_2$ is in $E$, while $v_1 v_3$ and $u_1 u_3$ are not in $D$ and $E$ respectively.

Any two complete symmetric digraphs with the same number of points are necessarily isomorphic. Thus we say that there is exactly one complete symmetric digraph of five points, since any two such digraphs are isomorphic, and we speak of it as the complete symmetric digraph of five points. Similarly, we may refer to the totally disconnected digraph of four points in which every point is an isolate, etc.

It is easy to show that the relation of isomorphism between pairs of digraphs has the properties of reflexivity, symmetry, and transitivity and is therefore an equivalence relation. To do this, one verifies that any digraph $D$ is isomorphic to itself; if $D_1$ is isomorphic to $D_2$, then $D_2$ is isomorphic to $D_1$; and if $D_1$ is isomorphic to $D_2$ and $D_2$ is isomorphic to $D_3$, then $D_1$ is isomorphic to $D_3$. 

![Figure 1.14](image-url)
All the digraphs with three points are shown in Figure 1.15. No matter what digraph with three points anyone will ever construct, it must be isomorphic to exactly one of these. Thus, we may say that there are sixteen digraphs with three points, or that the number of digraphs with three points and two lines is four, and so on.\footnote{Mandler and Cowan (1962) use the digraphs shown in Figure 1.15 to generate items in a learning experiment.}

\footnote{A formula for the number of digraphs with a given number of points and lines appears in Harary (1955b).}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{digraphs.png}
\caption{Figure 1.15}
\end{figure}

We now show how the isomorphism of any two digraphs may be determined by their adjacency matrices. Let $D_1$ and $D_2$ be digraphs whose adjacency matrices are $A_1$ and $A_2$ respectively. From matrix theory, the statement that \textit{two matrices are equal} means that they have the same size (i.e., the same number of rows and of columns) and for every $i$ and $j$, the $i,j$ entry of one and the $i,j$ entry of the other are equal.

\textbf{Theorem 1.2.} Two digraphs $D_1$ and $D_2$ are isomorphic if and only if for any ordering of the points of $D_1$, there is an ordering of the points of $D_2$ such that their adjacency matrices are equal, that is, $A_1 = A_2$.

This theorem may be proved in the following way. If there is an isomorphism between $D_1$ and $D_2$, take any ordering of the points of
Isomorphism of Digraphs

$D_1$ and reorder the points of $D_2$ in accordance with this isomorphism. Since by definition isomorphism entails the preservation of lines, there is a line from the $i$th point of $D_1$ to its $j$th point if and only if there is also a line from the $i$th point of $D_2$ to its $j$th point. This is guaranteed by the fact that the points of $D_2$ have been ordered in accordance with the isomorphism. Translated into matrix terms, this means that the $i, j$ entry of $A_1$ and the $i, j$ entry of $A_2$ must both be $1$ or both be $0$. Then $A_1 = A_2$ by the definition of equality of matrices.

Conversely, if $A_1 = A_2$, then clearly $D_1$ and $D_2$ are isomorphic.

If digraphs $D_1$ and $D_2$ are isomorphic, then any transmitter of $D_1$ corresponds to a transmitter of $D_2$. Similarly any receiver of $D_1$ corresponds to a receiver of $D_2$, and so forth. Since, by Theorem 1.2, two isomorphic digraphs have equal adjacency matrices, it follows that the sums of corresponding rows and columns of the two matrices are equal, and consequently the classification of any two corresponding points are the same.

Theorem 1.2 may be illustrated by the following example. Suppose that we are interested in three persons $u, v,$ and $w$, who display liking relationships as shown by digraph $D_1$ of Figure 1.16. As an empirical axiom, assume that whenever one person likes a second person, then the second person can influence the first. From this assumption, we obtain an influence structure for persons $u, v,$ and $w$, as represented by digraph $D_2$ of Figure 1.16. Clearly, we may order the points of each digraph so as to obtain the adjacency matrices $A(D_1)$ and $A(D_2)$, and it is evident that these two matrices are equal. By Theorem 1.2, $D_1$ and $D_2$ are isomorphic. This means that the liking and the influence
relationships among these three people have the same structure. And, both digraphs have one transmitter, one receiver, and one carrier. It should be noted, however, that the transmitter in the liking structure is the receiver in the influence structure, and conversely. Thus, if two isomorphic digraphs are interpreted as two different empirical relations on the same set of empirical entities, it does not necessarily follow that a given entity will correspond to a point of the same class in the two digraphs.

STRUCTURAL MODELS

We noted earlier that the social scientist may hope to gain certain benefits from employing digraph theory as a mathematical model. This possibility arises from the fact that he is interested in such empirical structures as groups of people, organizations, networks of communication, and systems of attitudes, beliefs, or alternatives, while digraph theory deals with structural properties in the abstract. As we have seen, the basic terms of digraph theory are point and line. Thus, if an appropriate coordination is made so that each entity of an empirical system is identified with a point and each relationship is identified with a line, then for all true statements about structural properties of the obtained digraph there are corresponding true statements about structural properties of the empirical system. The availability of a formal theory of structure, therefore, provides the researcher with conceptual tools for his analysis of empirical structures.

What empirical terms, or concepts of social science theory, can properly be coordinated to point and line? Consider first the term, point. In order to give an empirical interpretation of point, it is necessary to decide upon the appropriate unit of empirical data. Although the possibilities are almost limitless, five classes of entities may be suggested: persons, objects, places, events, and propositions. Entities from any of these classes may be identified with points. Thus, for example, the set of points $V$ of a particular digraph $D$ may correspond to the students enrolled in a course, the commodities in a market, the positions (offices) of an organization, the alternatives in a decision situation, or the propositions of an argument.

Suppose, then, that we have selected a set of empirical entities and have identified them with a set $V$ of points. In order to obtain a digraph $D$, we must have some way of identifying its set $X$ of lines. As we have seen, every line $x$ consists of an ordered pair $(u, v)$ of points, $fx = u$ and $vx = v$. An empirical relation on a set of people, such as liking, consists of a set of ordered pairs of people such that $(u, v)$ is in the relation if
and only if person \( u \) likes person \( v \). Thus, if we have sufficient information about the liking relation on a particular group of people, we may construct a digraph \( D \) of this empirical relation. Then, the presence of a line \( uv \) in \( D \) means that the person corresponding to point \( u \) likes the person corresponding to point \( v \), and the absence of line \( uv \) means that person \( u \) does not like \( v \). It is clear that other empirical relations on the same set of people may be treated in a similar fashion.

Essentially the same procedure may also be used to construct digraphs when points and lines are given other interpretations. Thus, the digraph of the competition structure of a market may be constructed by letting points stand for commodities and a line \( uv \) indicate that commodity \( u \) competes with commodity \( v \). The digraph of the authority structure of an organization may be constructed so that each point corresponds to a position and each line \( uv \) means that position \( u \) can exercise authority over position \( v \). In a realization where points stand for alternative outcomes, a line \( uv \) may represent the fact that alternative \( u \) is preferred to alternative \( v \). And, if points represent propositions, then line \( uv \) may indicate that proposition \( u \) implies proposition \( v \).

The empirical scientist clearly has great latitude in what empirical phenomena he coordinates to the primitives of digraph theory. Just how he proceeds will depend in part upon the nature of the phenomena he is studying. If his interests are in the social behavior of individuals, he may want to interpret points as persons and lines as interpersonal relationships of a type that interests him. Obviously, too, he will want to select an interpretation which will help him reach conclusions of significance for some theory about the world of empirical data. While we shall not attempt here to specify which interpretations should be made in order to achieve theoretical significance, it may be useful to discuss briefly some of the alternatives that must be faced in deciding upon a particular interpretation.

In interpreting point and line, the social scientist may be interested in relationships that indicate the possibility, the necessity, or the actual occurrence of an event. These different kinds of criteria are reflected by the words "can," "must," and "does," that is, by the use of the potential, imperative, and indicative moods. Thus, a line \( uv \) may be interpreted as "person \( u \) can communicate to person \( v \)," "\( u \) must communicate to \( v \)," or "\( u \) does communicate to \( v \)." There is nothing in the nature of digraph theory which requires that the interpretation refer to any particular modality.

Consider a simple digraph consisting of points \( v_1, v_2, v_3 \) and two lines \( v_1v_2 \) and \( v_2v_3 \). If the interpretation of a line is "can communicate message \( M \)," it follows that message \( M \) can go from \( v_1 \) to \( v_3 \) but not
that it will necessarily do so. If the interpretation is "must communicate," then if message \( M \) originates at \( v_1 \) it must reach \( v_r \). And if the interpretation is "does communicate," the digraph describes the actual transmission of the message. In most social psychological research on communication networks the interpretation has been "can communicate." And in this research it has become apparent that, in order to be able to predict completely the actual flow of communication, we must state certain "operating procedures" which people employ when working in any given network of potential communication.\(^{11}\)

Considerations of this sort suggest that a distinction should be made between descriptive and predictive interpretations of digraphs. The nature of this distinction may be illustrated by an example. Suppose that we have obtained friendship choices from a group of people. We can then describe the "friendship structure" of this group in terms of digraph theory by interpreting points as persons and lines as friendship choices. The concepts of digraph theory will apply to this structure, and its theorems will give factual information about this or any other structure. For example, a theorem to be presented in Chapter 12 asserts that if every point of a digraph \( D \) has outdegree 1, then \( D \) must have one of a very limited number of configurations. In fact, it follows from this theorem that every such digraph with four points must be isomorphic with one of those shown in Figure 1.17. We may conclude, then, that in a group of four people in which each person has exactly one friend, the only possible friendship structures are those shown in Figure 1.17. Factual information of this sort is of real value, but it does not lead to further predictions about the behavior of people unless

\(^{11}\) See, for example, Guetzkow (1960).
it is taken as an empirical axiom that "friendship choice" implies some particular resultant behavior. If, to achieve a predictive interpretation, an empirical axiom is set up so that "u chooses v" implies that u and v will tend to engage in social activities together, then we may draw the empirical conclusion that in a group of four people each of whom has exactly one friend, all will tend to engage in social activities together if and only if there are not two pairs of mutual friends. Although this conclusion appears obvious in such a small group, more complex conclusions may be derived in the same manner for groups of any size.

We see, then, that digraph theory will be useful to the social scientist in his efforts to describe the structural properties of empirical phenomena but that digraph theory, in and of itself, is not sufficient to derive empirical tendencies or laws from structural properties. Throughout this volume, in discussing the strictly structural properties of digraphs, we suggest a few rather simple empirical axioms which, when combined with digraph theory, result in certain empirical consequences. These are intended merely to illustrate how, in principle, empirical consequences may be derived. The reader will wish to consider for himself the consequences which he may deduce from other empirical axioms of interest to him.

**SUMMARY**

In this chapter we have been concerned with matters lying at the foundation of digraph theory. We began by considering an axiom system for the very general theory of nets. This system has four primitive terms: \( P_1 \): a set \( V \) of points, \( P_2 \): a set \( X \) of lines, \( P_3 \): a function \( f \) whose domain is \( X \) and whose range is contained in \( V \), and \( P_4 \): a function \( s \) whose domain is \( X \) and whose range is contained in \( V \). The system has two axioms: \( A_1 \): the set \( V \) is finite and not empty, and \( A_2 \): the set \( X \) is finite.

We next showed that an axiom system for relations is obtained simply by adding the axiom, \( A_3 \): no two distinct lines are parallel. Certain properties of relations were then defined and illustrated, and we saw how different kinds of orders on a set of points arise from various combinations of these properties.

With this background, we were able to place digraph theory in its more general mathematical context. The axiom system for digraph theory consists of that for relations with one additional axiom, \( A_4 \): There are no loops. Thus, a digraph is an irreflexive relation. We then saw how various properties of relations may be employed to generate
different kinds of digraphs. These will be encountered throughout the remainder of the book.

Many of the properties of digraphs are more conveniently handled in terms of matrices. In particular, there is a correspondence between a digraph and its adjacency matrix which allows a classification of points in terms of outdegree and indegree. It was shown that the sum of the outdegrees of all points, the sum of their indegrees, and the number of lines of a digraph are all equal. A detailed discussion of digraphs and matrices is given in Chapter 5.

Next, we considered the fundamental concept of isomorphism of two digraphs. With this concept we were able to give precise meaning to the notion that two empirical systems have the same structure. A theorem was then given, showing how the determination of the isomorphism of two digraphs can be obtained by means of their adjacency matrices.

At the beginning and the end of the chapter we discussed how the empirical scientist can use digraph theory. We pointed out that digraph theory is concerned with structural properties of sets of abstract elements called points and lines, whereas the empirical scientist is interested in empirical structures made up of empirical entities and relationships. If an appropriate coordination is made so that each empirical entity is identified with a point and each empirical relationship is identified with a line and if this is done in such a way that the axioms of digraph theory become true statements about the empirical world, then all true statements of digraph theory correspond to true statements about the empirical phenomena. When these requirements are met, the scientist may use the results of digraph theory in his treatment of empirical data.

It was noted that the empirical scientist also needs empirical axioms, in addition to those of digraph theory, in order to derive empirical tendencies or laws from structural properties.

In the next chapter we take up the fundamental properties of joining and reaching within digraphs.

EXERCISES

1. Among the digraphs of Figure 1.15, how many are symmetric? Asymmetric? Transitive? Complete? Disconnected?
2. Draw a digraph that contains an ordinary point and has a minimum number of points and lines. Draw another such digraph. Are there any others?
3. Draw a transitive digraph of three points and three lines. What is the classification of each point? Can one construct a transitive digraph whose points are all carriers?
4. Construct a $4 \times 4$ matrix which corresponds to an irreflexive, asymmetric, transitive relation. Draw the corresponding digraph. What kind of empirical situation might this digraph represent?

5. Construct a digraph and its corresponding adjacency matrix for the children's game "Paper, Rock, and Scissors" in which paper defeats rock, rock defeats scissors, and scissors defeats paper. (a) Characterize the relation in terms of reflexivity, symmetry, and transitivity. (b) What is the indegree and out-degree of each point?

6. Draw a digraph containing six points and five lines such that all these lines constitute the outbundle at the first point $v_1$. Write its adjacency matrix.

7. Verify that the two digraphs of Figure 1.14 are isomorphic by writing their adjacency matrices in such a way that they are equal.

8. Show that the two digraphs of Figure 1.18 are isomorphic. (a) Assign orderings to the points of $D_1$ so that the two adjacency matrices are equal. (b) Can this be done in more than one way?

9. Let an asymmetric digraph $D$ have three points and three lines. (a) Prove: If $D$ has exactly one carrier, then $D$ has one transmitter and one receiver. (b) Prove: If $D$ has one transmitter and one receiver, then $D$ is transitive. (c) Prove: If $D$ is transitive, then $D$ has exactly one carrier. (d) Construct an implication digraph for these three previous statements. (e) Use the fact that the relation of implication is transitive, and show from the implication digraph that the following statements are true: (1) If $D$ is transitive, then $D$ has one transmitter and one receiver. (2) If $D$ has exactly one carrier, then $D$ is transitive. (3) If $D$ has one transmitter and one receiver, then $D$ has exactly one carrier.

APPENDIX

In this book we make use of certain concepts of set theory. For the reader who wishes to refresh his memory of set-theoretic terminology and notation, we present the following very brief summary. Here, "set," "element," and "is an element of" are taken to be undefined terms with a natural intuitive meaning.
The universal set $U$ is the set of all elements under consideration. The empty set $\emptyset$ is the set which does not contain any elements. The set $B$ is a subset of a set $A$, written $B \subset A$, if every element of $B$ is in $A$. Two sets are called equal if each is a subset of the other. We say that $B$ is a proper subset of $A$ if $B$ is a nonempty subset of $A$ and $B$ does not equal $A$. Clearly every set is a subset of the universal set.

There are several important operations on sets. The union of two sets $A$ and $B$, written $A \cup B$, is the set consisting of all those elements which are in $A$ or $B$ (or both). The intersection of $A$ and $B$, denoted $A \cap B$, consists of those elements in both $A$ and $B$. If the intersection of $A$ and $B$ is empty, the sets are said to be disjoint. The difference $A - B$ contains all elements of $A$ which are not in $B$. The symmetric difference, written $A \oplus B$, is the set containing those elements in exactly one of the sets $A$ and $B$. Their symmetric difference can be thought of as the elements in $A$ or $B$ but not in both: $A \oplus B = (A \cup B) - (A \cap B)$. It can also be

**FIGURE 1.19**

![Diagram of set operations](image)
thought of as the elements of \( A \) not in \( B \) together with those in \( B \) not in \( A \): \( A \oplus B = (A - B) \cup (B - A) \), whence the name "symmetric difference." The complement of a set \( A \), denoted \( \bar{A} \), consists of the elements of \( U \) not in \( A \): \( \bar{A} = U - A \). These operations and their resulting sets are illustrated in Figure 1.19 by what are usually called Venn diagrams, in which the universal set \( U \) is a rectangular region of the plane and where the hatched area is the set resulting from the operation.

Some of the most useful set theoretic identities, whose validity may be readily verified by Venn diagrams, are as follows:

\[
\begin{align*}
A \cup U &= U & A \cap U &= A \\
A \cup \emptyset &= A & A \cap \emptyset &= \emptyset \\
A \cup \bar{A} &= U & A \cap \bar{A} &= \emptyset \\
A \cup A &= A & A \cap A &= A \\
A \cup B &= B \cup A & A \cap B &= B \cap A \\
A \cup (B \cap C) &= (A \cup B) \cup C & A \cap (B \cap C) &= (A \cap B) \cap C \\
A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) & A \cup (B \cap C) &= (A \cup B) \cap (A \cup C)
\end{align*}
\]
2 • Joining and Reaching

In every modern corporation
are channels of communication
along which lines, from foot to crown,
reports flow up and vetoes down.*
KENNETH E. BOULDING

Suppose that we are interested in studying the communication system
of an industrial organization and that we have data indicating for every
pair of people whether it is possible for each to give information about
some topic directly to another. From these data about the communication
system we may construct a digraph in which each point corresponds
to a person and each line uv indicates that person u can communicate
directly to person v. Given this digraph, we may ask whether information
possessed by one person (for example, a particular foreman) can reach
another person (for example, the president). If there is a line in the
digraph from the foreman to the president, the answer is, of course,
straightforward. But if the foreman is not adjacent to the president,
we need to examine the digraph further, for it is possible that information
can reach the president from the foreman indirectly through a
sequence of communications. Considerations of this sort make it clear
that two points u and v of a digraph D may be “joined” in some sense
even though D has no line uv.

If in our study of the communication system we find that information
can get in some way from person u to person v, we may regard v as
reachable from u. Considering the foreman and the president, we may

* Reprinted by permission from the March, 1958, issue of the Michigan Business Review,
published by the Graduate School of Business Administration, The University of
Michigan.
discover that neither can reach the other, only one can reach the other, or each can reach the other. It would not be at all surprising to find that the president can reach the foreman whereas the foreman cannot get through to the president. Clearly, then, reachability may, or may not, be symmetric.

There is another, and "weaker," way in which two persons u and v may be joined in a communication system. Suppose it is found that neither can reach the other but that each can reach a third person w. Since w is reachable from both u and v, we may regard u and v as joined, albeit weakly, by the communication system even though neither can get information to the other.

This brief discussion of the various ways in which two people may be joined serves to illustrate the need for carefully defined terms. In this chapter we consider concepts from digraph theory which are useful in making precise such intuitive notions as "joining" and "reaching." The fundamental concept for this purpose is semipath, a term defined broadly enough so as to be able to deal with the various kinds of joining suggested above. We shall see that there are two different kinds of semipaths and that the difference between them depends in a fundamental way upon the inclusion in them of the different classes of points discussed in the preceding chapter.

Closely related to the concept of joining is that of distance. If, for example, we know that two foremen can reach the president but that one can talk to him directly whereas the other must go through several people, we would regard the former as much "closer" to the president than the latter. By the same token, if the president can get information directly to the second foreman, we would say that the distance from the president to this foreman is less than the distance from this foreman to the president. In the following discussion we give precise meaning to the notion of the "distance" from one point to another in a digraph.

The material presented in this chapter is basic to all that follows in this volume. In particular, an understanding of the nature of semipaths will give the background needed for consideration, in the next chapter, of how the connectedness of complex structures may be characterized.

**SEMIPATHS AND PATHS**

A semipath joining \(v_1\) and \(v_n\) is a collection of distinct points, \(v_1, v_2, \ldots, v_n\) together with \(n - 1\) lines, one from each pair of lines, \(v_1v_2\), \(v_2v_3\), \(v_3v_4\), \(v_{n-1}v_n\) or \(v_nv_{n-1}\). Such a semipath is sometimes called a \(v_1 - v_n\) semipath.
A (directed) path from \( v_1 \) to \( v_n \) is a collection of distinct points, \( v_1, v_2, \ldots, v_n \), together with the lines \( v_1v_2, v_2v_3, \ldots, v_{n-1}v_n \) considered in the following order: \( v_1, v_1v_2, v_2v_3, \ldots, v_{n-1}v_n, v_n \).

In Figure 2.1 there is a path from \( v_1 \) to \( v_5 \). There is also a semipath joining \( v_1 \) and \( v_5 \). In fact, every path is a semipath. A semipath which is not a path is called a strict semipath. Accordingly, in Figure 2.1 there is a strict semipath joining \( v_1 \) and \( v_5 \). Clearly, there is no semipath joining \( v_3 \) and any other point of the digraph. If there is a path from \( u \) to \( v \), we say that \( v \) is reachable from \( u \). The following statements can therefore be made concerning the digraph of Figure 2.1: Points \( v_1 \) and \( v_4 \) are joined by a semipath but neither point is reachable from the other; \( v_1 \) and \( v_5 \) are joined, \( v_3 \) is reachable from \( v_1 \), but \( v_1 \) is not reachable from \( v_3 \); \( v_1 \) and \( v_5 \) are not joined and hence neither is reachable from the other.

The number of lines in a path is called its length. A geodesic from \( u \) to \( v \) is a path from \( u \) to \( v \) of minimum length. If there is a path from \( u \) to \( v \) in a digraph, then the distance from \( u \) to \( v \), denoted \( d(u, v) \), is the length of a geodesic from \( u \) to \( v \). It is important to note that in general the distance from \( u \) to \( v \) need not be equal to the distance from \( v \) to \( u \). For example, in Figure 2.2, \( d(v_1, v_2) = 1 \) and \( d(v_2, v_1) = 2 \). If there is no path from \( u \) to \( v \), the distance from \( u \) to \( v \) is called infinite, for there is no way to get from \( u \) to \( v \) in a finite number of steps. Symbolically, we write \( d(u, v) = \infty \).

For completeness, we note that a single point \( v \) is a path of length 0, called a trivial path. This is consistent with the definition of path for the following reasons: the single point \( v \) is a collection of distinct points and directed lines which begins and ends with a point and has the property (satisfied vacuously) that every line in this collection is preceded by its first point and followed by its second point. Since the length of a
path is the number of lines in it, the length of this path is 0. The distance from a point \( v \) to itself is, therefore, equal to zero: \( d(v, v) = 0 \).

The following theorem resembles the theorem of Euclidean geometry that the length of one side of a triangle is at most the sum of the lengths of the other two sides. It is known as the Triangle Inequality.

**Theorem 2.1.** If in a digraph, \( v \) is reachable from \( u \) and \( w \) is reachable from \( v \), then \( d(u, w) \leq d(u, v) + d(v, w) \).

By the hypothesis of this theorem, there are paths from \( u \) to \( v \) and from \( v \) to \( w \). By following these paths in succession, one gets from \( u \) to \( w \) in \( d(u, v) + d(v, w) \) steps. But since the distance \( d(u, w) \) is the fewest number of steps required to get from \( u \) to \( w \), \( d(u, w) \leq d(u, v) + d(v, w) \).

A subpath of a path \( L \) is a path contained in \( L \). Clearly, any subpath of a geodesic is itself a geodesic from its first point to its last point.

One may often want to find a geodesic from one point to another. If the digraph is small, this is easily done by inspection. However, in larger digraphs the following procedure may be helpful. Suppose we wish to find a geodesic from \( u \) to \( v \) in the digraph of Figure 2.3. First,

![Figure 2.3](image)

we assign to \( u \) the number 0 and to each point adjacent from \( u \) the number 1. Then the number 2 is assigned to every unnumbered point adjacent from a point marked 1, thereby designating the set of points at distance 2 from \( u \). If we continue in this manner, point \( v \) is eventually assigned a number which is its distance from \( u \). The numbers shown in Figure 2.3 are obtained in this way, so we know that \( d(u, v) = 4 \).

Hence a geodesic from \( u \) to \( v \) will have four lines. Let \( v_1 \) be any point at distance 3 from \( u \) such that \( v_1, v \) is a line of \( D \). Let \( v_2 \) be a point at
distance 2 from $u$ such that $v_2v_3$ is in $D$, and let $v_i$ be any point such that both lines $uv_i$ and $v_1v_2$ are in $D$. Then we have a geodesic from $u$ to $v$, whose points are $u, v_1, v_2, v_3, v_i$.

In our example, there is only one geodesic from $u$ to $v$. Of course, it may happen that there is more than one geodesic from one point to another, as for example from $u$ to $w$ in the digraph of Figure 2.3. However, the procedure is still applicable, and if we take all feasible choices for $v_3, v_2,$ and $v_1,$ we shall find all geodesics from $u$ to $w$. The procedure is stated in general terms in the following theorem.

**Theorem 2.2.** (Geodesic algorithm.)¹ Let $v$ be reachable from $u$ with $d(u, v) = n$. The following algorithm can be used to construct any geodesic from $u$ to $v$. First, find a point $v_{n-1}$ such that $d(u, v_{n-1}) = n - 1$ and line $v_{n-1}v$ is in $D$. Then find a point $v_{n-2}$ such that $d(u, v_{n-2}) = n_1 - 2$ and $v_{n-2}v_{n-1}$ is in $D$. Continue this process until finding a point $v_1$ such that $d(u, v_1) = 1$ and $v_1v_2$ is in $D$. Then the path $uv_1v_2v_{n-1}v$ is a geodesic from $u$ to $v$.

In order to gain a deeper understanding of the nature of semipaths, let us examine more closely how semipaths are constituted. For this purpose the classification of points presented in Chapter 1 will be employed. The semipath in Figure 2.4(a) is, of course, the primordial building block of any digraph, for it consists of a single directed line. If we construct semipaths containing exactly two lines, we find that the only possible configurations are those shown in Figure 2.4(b), (c), and (d). The letters $t, r, c$, and $l$, indicated at the points, stand for transmitter, receiver, and carrier respectively. It is clear that the sequences of letters $tcr, rtr$, and $trt$, indicating the classification of successive points, contain all the information required to draw each of these semipaths. It must be understood that when any of these sequences is "spelled backwards," the result still determines exactly the same semipath. Thus, both $tcr$ and $cnt$ determine the semipath of Figure 2.4(b). These considerations lead to the following conclusion.

**Theorem 2.3.** A sequence of letters consisting of $c$'s, $r$'s, and $t$'s describes a nontrivial semipath if and only if neither the first nor the last is $c$, there is an $r$ between any two occurrences of $t$'s, and there is a $t$ between any two $r$'s.

As an illustration of this theorem we note that the $v_1v_2v_3$ semipath in the digraph of Figure 2.1 can be described by $tcr$ or by $tcr$.² Between the two occurrences of $t$ there is an $r$.

¹ This algorithm is due to Moore (1959).
² Since the "word" describing a semipath of length 3 contains four letters, we have carefully selected our alphabet so as to avoid any possible embarrassment.
Corollary 2.3a. A nontrivial semipath is a path if and only if it can be described with a sequence of letters containing exactly one $t$ and one $r$. It is a strict semipath if and only if its description requires at least two $r$'s or at least two $r$'s.

We see that the character of a semipath is to a large extent determined by its $r$-points and $t$-points. It is often convenient therefore to suppress the $c$-points. The suppression of the carriers of a semipath $L$ corresponds to "erasing" all $c$'s from the sequence of letters describing $L$. The semipath resulting from the suppression of all carriers of $L$ is called its suppressed semipath. Thus, for example, the digraph of Figure 2.4(a) is the suppressed semipath of the digraph of Figure 2.4(b). The next statement follows immediately from Theorem 2.3 and the definition of suppression.

\[ \begin{array}{cc}
\text{\begin{tikzpicture}[->,>=stealth',auto, scale=0.7]
  \node (n1) at (0,0) {t};
  \node (n2) at (1,0) {r};
  \path (n1) edge (n2);
\end{tikzpicture} & \begin{tikzpicture}[->,>=stealth',auto, scale=0.7]
  \node (n1) at (0,0) {r};
  \node (n2) at (1,0) {t};
  \path (n1) edge (n2);
\end{tikzpicture}} \\
\text{(a)} & \text{(c)} \\
\text{\begin{tikzpicture}[->,>=stealth',auto, scale=0.7]
  \node (n1) at (0,0) {t};
  \node (n2) at (1,0) {c};
  \node (n3) at (2,0) {r};
  \path (n1) edge (n2) (n2) edge (n3);
\end{tikzpicture} & \begin{tikzpicture}[->,>=stealth',auto, scale=0.7]
  \node (n1) at (0,0) {r};
  \node (n2) at (1,0) {c};
  \node (n3) at (2,0) {t};
  \path (n1) edge (n2) (n2) edge (n3);
\end{tikzpicture}} \\
\text{(b)} & \text{(d)} \\
\end{array} \]

**FIGURE 2.4**

Corollary 2.3b. The suppression of the carriers in a semipath results in a semipath consisting entirely of $t$-points and $r$-points in alternation, and the numbers of $t$-points and $r$-points are unchanged by suppression.

A maximal path in a semipath $L$ is a path contained in $L$ but is not a subpath of any longer path in $L$. A maximal path in a semipath must have one $r$-point and one $t$-point of the given semipath; its other points are all $c$-points. Clearly, each line of a suppressed semipath corresponds to a maximal path in the original semipath. Two maximal paths in a semipath $L$ can have but one point in common, which must be a $t$-point of both or an $r$-point of both, since it must be a $t$-point or an $r$-point of $L$. Such a point common to two paths in $L$ is called a linking point. Thus in any strict semipath there are at least two maximal paths and at least one linking point.

A semipath can be regarded as built up from its maximal paths. If $L$ is a semipath joining $u$ and $v$, then it can be thought of as constructed from maximal paths $L_1$, $L_2$, ..., $L_n$ and linking points $v_1$, $v_2$, ..., $v_{n-1}$, where $L_1$ joins $u$ and $v_1$, $L_2$ joins $v_1$ and $v_2$, ..., $L_n$ joins $v_{n-1}$ and $v$. 
The points $u, v_1, v_2, \ldots, v_{m-1}, v$ are alternately receivers and transmitters. Only the linking points lie on more than one maximal path, and linking point $v_i$ lies only on $L_i$ and $L_{i+1}$. We note that the number of maximal paths exceeds the number of linking points by 1. The following corollary is a particular case of some of these remarks.

**Corollary 2.3c.** If two transmitters are joined by a semipath $L$ that contains no other transmitter, then $L$ has exactly one receiver. Furthermore, for any point $v$ in $L$ there is a path from $v$ to the receiver.

**DIRECTIONAL DUALITY**

In this section we study the operation of taking the "converse" of any digraph. We shall see that this operation, which involves reversing the direction of every line of a given digraph, sets the stage for a powerful principle called "directional duality." This principle will enable us to establish certain theorems without effort once we have proved other corresponding theorems.

Note, first, that the digraphs of Figures 2.4(c) and (d) are related to each other in a particular way: either one can be obtained from the other simply by reversing the directions of all lines. Given a digraph $D$, its *converse* $D'$ is the digraph with the same set of points such that for any two points $u$ and $v$ the line $uv$ is in $D'$ if and only if $vu$ is in $D$. The digraphs of Figures 2.4(c) and (d) are converses of each other, and the digraph of Figure 2.2 is isomorphic to its own converse.

**Theorem 2.4.** The converse of the converse of a digraph $D$ is $D$ itself; symbolically, $D'' = D$.

To prove this statement observe that $D$ and $D'$ have the same set of points. We must show that they have the same lines. By the definition of $D'$, $uv$ is a line of $D'$ if and only if $vu$ is a line of $D$. Since $D''$ is the converse of $D'$, $uw$ is a line of $D''$ if and only if $uw$ is a line of $D$. Thus $uv$ is in $D''$ if and only if it is in $D$. Therefore, they have the same lines.

\[ \text{FIGURE 2.5} \]

\[ ^3 \text{The approach employed here is that presented by Harary (1957).} \]
The two digraphs shown in Figure 2.5 serve to illustrate this theorem. Note that the digraph of Figure 2.5(b) is the converse of that in Figure 2.5(a). In addition, the digraph of Figure 2.5(a) is the converse of the one in Figure 2.5(b). Thus, in accordance with Theorem 2.4, the converse of the converse of the digraph in Figure 2.5(a) is itself the digraph in Figure 2.5(a).

A dual operation means an operation on digraphs which, when applied twice, results in the original digraph. Thus, Theorem 2.4 tells us that one dual operation is that of taking converses. For each dual operation there is an associated collection of dual concepts; in the particular case of the converse operation, these are called converse concepts. More specifically, the converse of a concept concerning digraphs is one which results in place of the concept, when the operation of converse is applied to a digraph. The following theorem specifies some of the more frequently occurring converse concepts. In each case, the words before the verb "becomes" refer to the original digraph $D$, while the words after "becomes" refer to the converse digraph $D'$. Each of the statements in the theorem follows at once from the definition of the operation of converse.

**Theorem 2.5.** The following ten statements specify what happens to certain parts of a digraph $D$ when the converse digraph $D'$ is formed.

1. Every point $v$ of $D$ becomes exactly the same point $v$ in $D'$.
2. Every line $uv$ becomes line $vu$; every path from $u$ to $v$ becomes a path from $v$ to $u$.
3. Every strict semipath joining $u$ and $v$ becomes a strict semipath joining these same two points.
4. Every transmitter becomes a receiver.
5. Every receiver becomes a transmitter.
6. Every isolate becomes an isolate.
7. Every carrier becomes a carrier.
8. Every ordinary point becomes an ordinary point.
9. The indegree of $v$, $\text{id}(v)$, becomes the outdegree of $v$, $\text{od}(v)$.
10. The outdegree of $v$ becomes the indegree of $v$.

A concept is self-dual with respect to a dual operation if it remains the same after the operation is performed on the digraph. We observe that certain concepts, such as point, isolate, and carrier, are self-dual directionally, that is, they remain the same under the operation of taking the converse of a digraph.

The dual of a statement about a digraph refers to that statement obtained upon replacing each concept in the statement by the dual
concept. For example, the directional dual of the statement that $v$ is a transmitter of a digraph is that $v$ is a receiver, and the directional dual of the statement that there is a path from $u$ to $v$ is that there is a path from $v$ to $u$. In both of these examples, we see that in forming the directional dual of a statement about a digraph, we replace each concept by its converse concept, that is, its directionally dual concept.\footnote{When he married, he jumped into the frying pan out of the fire.}

In general, any duality principle has the following two properties:
(a) The dual of the dual of a statement is the original statement. (b) The dual of a true statement is true. We now state and prove the directional duality principle.

**Directional Duality Principle.** For each theorem about digraphs, there is a corresponding theorem obtained by replacing every concept by its converse concept.

In order to prove this principle we consider the nature of a theorem about digraphs. Ultimately, any such theorem can be recast in the following form: For any digraph $D$, if $P$ is a true statement, then $Q$ is a true statement, that is, $P$ implies $Q$. Let us denote the respective dual statements by $P'$ and $Q'$. The directional duality principle asserts that if the statement “$P$ implies $Q$” holds for any digraph $D$, then the statement “$P'$ implies $Q'$” also holds for any digraph $D'$. We begin the proof by taking as a given theorem, whose proof is already known, the following statement:

1. For any digraph $D$, if $P$ is true then $Q$ is true.

   It is obvious that if $P$ is true for $D$, then $P'$ holds for $D'$. Likewise if $P'$ holds for $D'$, then $P$ holds for $D$. Obviously $P' = P$, and by Theorem 2.5, $D' = D$. Thus, $P$ holds for $D'$ if and only if $P'$ holds for $D'$; and of course, the same is true for statement $Q$. Statement (1) can be reworded, therefore, in a logically equivalent way:

2. For any digraph $D$, if $P'$ is true for $D'$ then $Q'$ is true for $D'$.

   The preceding statement can now be stated a bit more briefly by not mentioning digraphs $D$ explicitly:

3. For any digraph $D'$, if $P'$ is true then $Q'$ is true.

   Thus, statement (1) is equivalent to saying that for any digraph which is the converse of some digraph, if $P'$ is true then $Q'$ is true. But every digraph is the converse of some digraph; therefore, we have the following logically equivalent form of statement (3):

4. For any digraph $D$, if $P'$ is true then $Q'$ is true.
This last statement (4) is the theorem corresponding to statement (1) obtained by replacing every concept in the original theorem by its converse concept. Since each of the statements (2)-(4) is logically equivalent to the preceding one, it follows that statements (4) and (1) are equivalent to each other. But statement (1) is the given theorem which is known to be true; hence statement (4) is also true and the directional duality principle is proved.

The way in which the directional duality principle is employed in digraph theory may be illustrated by reference to Corollary 2.3c. This corollary asserts that any semipath $L$ with exactly two transmitters which joins these two points has exactly one receiver, and that for any point $v$ in $L$ there is a path from $v$ to the receiver. On interchanging the words "to" and "from," and the words "transmitter" and "receiver," we obtain the following dual corollary.

**Corollary 2.3c'.** If two receivers are joined by a semipath $L$ that contains no other receiver, then $L$ has exactly one transmitter. Furthermore, for any point $v$ in $L$ there is a path to $v$ from the transmitter.

It may happen that the directional dual of a statement is the very same statement. If this is true for a theorem, we say that the theorem is self-dual. Consider, for example Theorem 2.3, which asserts that a sequence of letters consisting of $c$'s, $r$'s, and $t$'s describes a nontrivial semipath if and only if neither the first nor the last is $c$, there is an $r$ between any two occurrences of $r$'s, and there is a $t$ between any two $r$'s. Strictly speaking, the directional dual of this theorem is the following statement: A sequence of letters consisting of $c$'s, $r$'s, and $t$'s describes a nontrivial semipath if and only if neither the first nor the last is $c$, there is a $t$ between any two occurrences of $r$'s, and there is an $r$ between any two $r$'s. Even though the order of stating $r$'s and $t$'s is changed in this converse statement, the logical meaning of the two statements is exactly the same, and Theorem 2.3 is therefore self-dual.

**THE JOINING OF PAIRS OF POINTS**

In this section we investigate various ways in which pairs of points can be joined in a digraph. The analysis of joining relations is fundamental to much that follows in this book. Joining relations will be used in Chapter 3 to define the kinds of connectedness which digraphs may display. In order to make this analysis we need to generalize the concepts of path and semipath to allow repetition of their points and lines.

A (point-line) sequence is an alternating sequence of points and lines which begins and ends with a point and has the property that each line
is preceded by its first point and followed by its second point. Thus a sequence may be written in the form: \( v_1, v_2, v_3, \ldots, v_{n-1}, v_n \). This sequence is determined by its points \( v_1, v_2, \ldots, v_n \) or by its lines \( v_1v_2, v_2v_3, \ldots, v_{n-1}v_n \). For brevity it is sometimes indicated by \( v_1v_2 \ldots v_n \). The point \( v_1 \) is the initial point of this sequence; \( v_n \) is the terminal point.

We say that this is a sequence from \( v_1 \) to \( v_n \). There is no restriction on the number of times a point or line may be repeated in the same sequence. Thus, even though we have used different symbols (given by subscripts) for the points, it is not necessary that these indicate distinct points. A sequence is called open if \( v_1 \) and \( v_n \) are distinct points. The difference between an open sequence and a path is that in a path all points and lines must be distinct.

**Theorem 2.6.** A sequence from \( u \) to \( v \) contains a path from \( u \) to \( v \).

To prove this theorem, let \( L \) be a sequence from \( u \) to \( v \) and let \( w \) be any point of \( L \). If \( w \) occurs more than once in \( L \), remove from \( L \) all points and lines between the first and last occurrence of \( w \) and also remove this last occurrence of \( w \). The result of repeated applications of this procedure is a path from \( u \) to \( v \) contained in \( L \).

**Corollary 2.6a.** If there is a sequence from \( u \) to \( v \), then \( v \) is reachable from \( u \).

These concepts may be illustrated by reference to Figure 2.6. Consider the sequence \( L \) from \( v_1 \) to \( v_5 \):

\[
L = v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_3, v_6, v_4, v_7, v_2, v_1
\]

By application of the procedure described in the proof of Theorem 2.6, we see that there is a path \( v_1v_7v_3 \) from \( v_1 \) to \( v_3 \). And by Corollary 2.6a, \( v_3 \) is reachable from \( v_1 \). It is also evident that there is no sequence from \( v_1 \) to \( v_6 \). Thus, \( v_6 \) is not reachable from \( v_1 \).

If the initial and terminal points of a sequence are the same point, the sequence is said to be closed. If the sequence \( L \) from \( u \) to \( v \) in Theorem 2.6 is closed, then \( u = v \) and the construction in the proof of this theorem leaves the single point \( u \), which as we have seen is a path of length 0.

Another example of a closed sequence is the following one of the digraph

---

**Figure 2.6**
in Figure 2.6: \( v_3, v_2, v_5, v_4, v_3, v_2, v_4, v_5, v_3, v_2 \). This sequence has \( v_2 \) as both its initial and terminal points and is therefore closed.

A cycle is obtained from a nontrivial path on adding a line from the terminal point to the initial point of the path. For example, consider in Figure 2.6 the path \( v_2, v_5, v_4 \). Upon adding the line \( v_4, v_2 \) to this path, we obtain the cycle \( v_2, v_5, v_4, v_2 \). This cycle is precisely the same as the closed sequence of the preceding paragraph.

We note that a single point \( v \) is a (trivial) closed sequence, for it begins and ends with \( v \) and vacuously satisfies the definitional requirement concerning lines.

The next theorem states a result for closed sequences which is analogous to Theorem 2.6.

**Theorem 2.7.** If \( v \) is a point of a nontrivial closed sequence \( L \), then \( v \) is in a cycle contained in \( L \).

A semisequence is an alternating sequence of points and lines which begins and ends with a point and has the property that each line is incident with the point before it and the point after it. Unlike a sequence, it is not necessary that two consecutive lines of a semisequence have consistent directions. Thus a semisequence joining \( v_1 \) to \( v_n \) may be written in the form: \( v_1, v_2, v_3, \ldots, v_{n-1}, v_n \). Note that neither the points nor the lines of a semisequence need be distinct. In the digraph of Figure 2.6 there is a semisequence beginning with \( v_1 \) and ending with \( v_n \), which consists of the sequence \( L \) from \( v_1 \) to \( v_4 \) listed above together with the line \( v_5, v_2 \) and the point \( v_5 \).

**Theorem 2.8.** A semisequence joining \( u \) and \( v \) contains a semipath joining them.

A semicycle is obtained from a semipath on adding a line joining the terminal point and the initial point of the semipath. Note that every cycle is a semicycle, since every path is a semipath. The digraph in Figure 2.7 contains three semicycles: \( Z_1 \), the semipath \( v_2, v_3, v_1 \); together with line \( v_1, v_2 \); \( Z_2 \), the semipath \( v_1, v_2, v_4 \); together with line \( v_4, v_1 \); and \( Z_3 \), the semipath \( v_1, v_2, v_4 \); together with line \( v_4, v_2 \). Clearly, \( Z_1 \) is not a cycle, whereas \( Z_2 \) and \( Z_3 \) are.

The length of a sequence or a semisequence is the number of occurrences of lines in it. Thus, in particular, the length of a path, cycle, semipath, or semicycle is the number of lines in it. If any of these contains all the points of the digraph \( D \), it is said to be complete. Thus the semisequence \( v_1, v_2, v_3, v_4, v_5, v_6 \) of the digraph of Figure 2.6 is complete and of length 6. There is no complete sequence, path, semipath, semicycle, or cycle in Figure 2.6.
We may now consider a method for describing different ways in which a pair of points may be joined in a digraph. We say, for convenience, that every pair of points \( u \) and \( v \) of any digraph are 0-joined. They are 1-joined if there is a semisequence joining them, 2-joined if there is a sequence from one to the other, and 3-joined if there is a sequence from each to the other. By using Theorems 2.6 and 2.8, we see at once that these definitions can be stated equivalently in terms of semipaths and paths.

In Figure 2.8 five digraphs are shown. These illustrate the different kinds of joining that pairs of points may have. The kinds of joining in each digraph are presented in Table 2.1, which is constructed in the following manner. Since, by definition, every pair of points of a digraph is 0-joined, we enter 0 for every pair in each digraph. Since there are no lines in digraph \( D_1 \), these are the only entries for this digraph. For any two points joined by a semipath, we enter 1, and for any two points joined by a path we enter 2. Thus, for example, we enter both 1 and 2 for each pair of points in \( D_4 \) and \( D_5 \), but we enter only 1 for the pair \( v_1, v_3 \) in \( D_3 \) since they are joined by a semipath but not a path. Finally, we enter 3 for every pair of points that are mutually reachable. This last condition is met only in \( D_4 \), which consists of a cycle and therefore has a path in each direction between every pair of points.

<table>
<thead>
<tr>
<th>Pairs of Points</th>
<th>( D_1 )</th>
<th>( D_2 )</th>
<th>( D_3 )</th>
<th>( D_4 )</th>
<th>( D_5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( v_1, v_2 )</td>
<td>0</td>
<td>0, 1, 2</td>
<td>0, 1, 2</td>
<td>0, 1, 2</td>
<td>0, 1, 2, 3</td>
</tr>
<tr>
<td>( v_1, v_3 )</td>
<td>0</td>
<td>0</td>
<td>0, 1</td>
<td>0, 1, 2</td>
<td>0, 1, 2, 3</td>
</tr>
<tr>
<td>( v_2, v_3 )</td>
<td>0</td>
<td>0</td>
<td>0, 1, 2</td>
<td>0, 1, 2</td>
<td>0, 1, 2, 3</td>
</tr>
</tbody>
</table>
It is evident from the definitions and from Table 2.1 that if two points are 3-joined they are also 2-joined, if they are 2-joined they are also 1-joined, and if they are 1-joined they are 0-joined. For this reason, we frequently indicate the kind of joining of a pair of points simply by writing the highest degree of joining.

REALIZATIONS OF REACHABILITY AND JOINING

The concepts of reachability and joining are so fundamental in digraph theory that it will be useful at this point in our study to examine them more closely.

Reachability

As we have seen, point \( v \) is reachable from \( u \) in a digraph \( D \) if and only if there is a sequence from \( u \) to \( v \). It will be fruitful to conceive of reachability as a relation on the set of points \( V \) of a digraph \( D \). Clearly, this is a different relation from that of adjacency, which corresponds to the set of lines of \( D \). What properties does the relation of reachability have? First, reachability is reflexive since every point of \( V \) is
reachable from itself by a path of length 0. Second, it is transitive, for if there is a sequence \( L_1 \) from \( u \) to \( v \) and a sequence \( L_2 \) from \( v \) to \( w \), then there is a sequence \( L_3 \) from \( u \) to \( w \), namely the union of \( L_1 \) and \( L_2 \). Finally, it need be neither symmetric nor asymmetric, for if there is a sequence from \( u \) to \( v \), there may, or may not, be a sequence from \( v \) to \( u \).

As mentioned in Chapter 1, a reflexive and transitive relation is known as a quasi-order. Thus, the relation of reachability is a quasi-order.

The properties of the reachability relation must be kept in mind when we consider possible realizations of digraphs. Let us illustrate what is involved by means of a communication network of a group of people.\(^5\)

This network may be represented by a digraph if we let each person of the group correspond to a point of a digraph \( D \) and let the relationship "\( u \) can transmit message \( M \) directly to \( v \)" correspond to the line \( uv \) of \( D \).

In order to draw conclusions about the flow of message \( M \) through this network, we need to make precise empirical assumptions about the way in which the network operates. For convenience, we state the following simple assumptions.

**E2.1.** A person may possess message \( M \) only by originating it or receiving it from another person in the group.

**E2.2.** If a person possesses message \( M \), he will transmit it to all persons possible (that is, all persons adjacent from him in the digraph of the network).

We see immediately that if person \( v_i \) possesses message \( M \) and there is a sequence in the digraph of the network of the form \( v_i v_2 \ldots v_n \), then message \( M \) will eventually reach every person in this sequence. Thus, the empirical relationship "message \( M \) will reach \( v \) from \( u \)" can be coordinated to the graphical relationship "\( v \) is reachable from \( u \)."

In general, we represent an empirical relation on a set of empirical entities by a digraph, where each relationship of the relation corresponds to a line. If we wish in the realization to make use of reachability in digraphs, we must be able to make the coordination between a sequence in \( D \) and some underlying relation in the empirical world.

Let us return now to empirical assumptions E2.1 and E2.2 to see how these, together with the concept of reachability, lead to certain empirical conclusions. The first conclusion merely makes explicit the consequences of E2.2 noted above.

---

\(^5\) The use of graphs to represent the communication structure of a group was first employed by Bavelas (1950). For a review of subsequent work, see Glimmer and Glaser (1961).
C2.1. If person \( v \) possesses \( M \), then \( M \) will reach everyone reachable from \( v \) in the digraph of the network.

The next two conclusions follow immediately from this, when taken together with the definitions of 2-joining and 3-joining.

C2.2. Let persons \( u \) and \( v \) be 2-joined in the digraph of a communication network. Then at least one of the following statements is true: (a) Person \( u \) is reachable from \( v \), and if \( v \) originates \( M \) it will reach \( u \). (b) Person \( v \) is reachable from \( u \), and if \( u \) originates \( M \) it will reach \( v \).

C2.3. If in the digraph of a communication network \( u \) and \( v \) are 3-joined, then \( u \) and \( v \) can engage in two-way communication; that is, if \( u \) originates \( M \) it will reach \( v \), and if \( v \) originates \( M \) it will reach \( u \).

From C2.1 and the definition of a closed sequence, we obtain the next conclusion.

C2.4. Let \( D \) be the digraph of a nontrivial communication network such that \( v \) lies in a closed sequence \( L \). Then, if \( v \) originates \( M \), he will also receive it from another person in \( L \).

Finally, with the aid of the definition of a complete sequence we obtain the next conclusion.

C2.5. If in the digraph of a communication network there is a complete sequence \( L \), then \( M \) will reach everyone in the network if it originates with the person corresponding to the initial point of \( L \).

It should be evident that analogous conclusions can be obtained from any empirical realization in which an empirical relationship can be coordinated to a line of a digraph and an underlying empirical relationship can be coordinated to the reachability of pairs of points.

**Joining**

We saw above that two points \( u \) and \( v \) are joined in a digraph if and only if there is a semisequence, or semipath, between them. The difference between reachability and joining is that in the latter instance we ignore the direction of lines. Just as in the case of reachability, we may conceive of joining as a relation on the set of points \( V \) of a digraph \( D \). Joining is reflexive, for every point of \( V \) is joined to itself by a semipath of length 0. It is transitive, for if there is a semipath \( L_1 \) between \( u \) and \( v \) and a semipath \( L_2 \) between \( v \) and \( w \), then there is a semipath \( L_3 \) between \( u \) and \( w \). But joining differs from reachability in that it is symmetric, for if a semipath joins \( u \) and \( v \), the same semipath obviously joins \( v \) and \( u \). From the theory of relations, we may conclude that joining is an equivalence relation, since it is reflexive, transitive, and symmetric.
These observations have implications for realizations that make use of the joining of points in a digraph. Consider, for example, the communication network shown in Figure 2.9. (Note that the entire figure is one digraph, that is, the group consists of six people.) We see that \( e_1 \) and \( e_3 \) are joined by a semipath, as are \( e_2 \) and \( e_6 \). However, \( e_1 \) and \( e_4 \) are not joined in the digraph. Even if we ignore the direction of the lines of this digraph, we observe that \( e_1 \) and \( e_3 \) are in the same piece of the network, \( e_2 \) and \( e_4 \) are in the same piece, but \( e_1 \) and \( e_4 \) are in different pieces. The underlying empirical relation which is expressed by joining, then, is "being in the same piece of the network." Clearly, this is an equivalence relation and appropriately coordinated to joining in a digraph.

In comparing the relations of reachability and joining, it appears that joining is the weaker one. This observation is reflected in the fact that the more interesting realizations are ordinarily concerned with reachability. Joining, nevertheless, is an important feature of digraphs, and we shall encounter it repeatedly throughout this book. In the next chapter, we shall develop in greater detail an empirical realization that makes use of joining.

**UNIPATHIC DIGRAPHS**

A digraph is called *unipathic* if, whenever \( v \) is reachable from \( u \), there is exactly one path from \( u \) to \( v \). Obviously, every path in a unipathic digraph is a geodesic. Clearly a digraph is unipathic if and only if it has no semicycle which is the union of two paths from one point to another. Although every digraph with no semicycles is unipathic, the converse is not true, as shown by a digraph consisting of a single cycle. We illustrate unipathic digraphs by Figure 2.10. In Figure 2.10(a) there are no semicycles at all. Figure 2.10(b) consists of a single semicycle which is, however, unipathic. In Figure 2.10(c) there are three cycles, but no two of these cycles have a line in common.
Theorem 2.9. If $D$ is unipathic, then no two cycles of $D$ have a common line.

Assume that there is a line $x$ in two different cycles $Z_1$ and $Z_2$ of the unipathic digraph $D$. Then there is a path in $D$ from $sx$ to $fx$ along $Z_1$ and another such path along $Z_2$, contrary to the hypothesis.

Theorem 2.10. If $D$ is unipathic and $Z$ is a cycle, there is no line in $D$ joining any two points of $Z$ which does not lie in $Z$.

If there is a line $x$ joining two points of $Z$ such that $x$ is not in $Z$, the line $x$ itself is one path from $fx$ to $sx$, and there is another path from $fx$ to $sx$ along the cycle $Z$, contradicting the hypothesis that $D$ is unipathic.

SUMMARY

In this chapter we have examined different ways in which two points $u$ and $v$ of a digraph may be joined. If there is a path from $u$ to $v$, then $v$ is reachable from $u$. If they are joined only by a strict semipath, neither is reachable from the other. We showed that a strict semipath
contains a series of maximal subpaths linked together by linking points, each of which is a transmitter or a receiver of the semipath.

After discussing the operation of taking the converse of a digraph, we provided a powerful tool known as the directional duality principle. With the aid of this principle new theorems and corollaries can be obtained immediately from proven theorems and corollaries.

We then defined four kinds of joining that pairs of points may display. Examples were provided to show how the concepts of this chapter may be employed in empirical realizations. It was seen that the graphical concepts of reachability and joining may be coordinated to underlying empirical relations once a primary coordination has been made for the concepts of point and line. In the next chapter we shall show how the concepts of this chapter may be employed to define different kinds of connectedness which more complex structures may possess.

EXERCISES

1. Consider the digraph $D$ in Figure 2.11.

![Figure 2.11](image)

(a) Write a sequence from $v_1$ to $v_3$ that is not a path. (b) How many paths are there from $v_3$ to $v_4$? From $v_4$ to $v_5$? (c) How many semipaths are there joining $v_1$ and $v_3$? (d) How many semicycles are there in $D$? (e) How many cycles? (f) Write a semisequence beginning with $v_1$ and ending at $v_3$ that is not a sequence or a semipath.

2. Consider the following propositions.

$p_1$: $L$ is a semisequence joining $v_1$ and $v_4$.
$p_2$: $L$ is a sequence from $v_1$ to $v_4$.
$p_3$: $L$ is a semipath joining $v_1$ and $v_4$.
$p_4$: $L$ is a path from $v_1$ to $v_4$.

(a) Verify that the digraph $D$ shown in Figure 2.12 is the total implication digraph for these propositions. (b) Why is there no line $p_1p_2$? $p_1p_3$? $p_2p_4$? (c) Why are there no lines joining $p_3$ and $p_4$?
3. Consider the digraph $D$ of Exercise 1. (a) Construct the converse digraph of $D$, $D' = E$. (b) Identify the paths from $v_1$ to $v_3$ in $D$; in $E$. (c) What is the value of $d(v_1, v_3)$ in $D'$? In $E'$? (d) What is the value of $d(v_3, v_2)$ in $D'$? In $E'$? (e) What is the digraph $E'$?

4. Prove that $D$ is symmetric if and only if $D$ and its converse $D'$ have exactly the same lines.

5. What is the directional dual of each of the following statements? (a) If every two points of a digraph are joined by exactly one line, then there is a point $v$ whose distance to any other point is at most 2. (b) If a semipath has the property that given any two of its points at least one is reachable from the other, then this semipath is a path.

6. Use Theorem 2.3 to show that a semipath is a path if and only if it contains exactly one maximal path.

The last four exercises involve the following terminology: Given a semipath $P$, let $P$ be the number of maximal paths in $L$, let $R$ and $T$ be the number of $r$-points and $t$-points, respectively. Let $R_1$ and $R_2$ be the number of $r$-points with total degree 1 and 2, respectively; and let $T_1$ and $T_2$ be defined similarly.

7. Show that $P = R + T - 1 = R_1 + T_1 + 1$.

8. Show that the following three statements are equivalent: $P$ is even; the difference between $R_1$ and $T_1$ is 1; and the difference between $R$ and $T$ is 1.

9. Show that the following statements are equivalent: $P$ is odd; $R_2 = T_2$; and $R = T$.

10. Let $L$ be a semipath for which $T_1 = 2$, $T_2 = 1$, and $R_2 = 2$. (a) What is the directional dual of this statement? (b) Find $P$, the number of maximal paths in $L$. (c) How many maximal paths does the directional dual of $L$ have? (d) What is the smallest number of points $L$ can have? Can $L$ have 12 points?