PSYCHOMETRIKA—VOL. 31, NO. 2
JUNE, 1966

THE NUMBER OF PATHS AND CYCLES IN A DIGRAPH*

DORWIN CARTWRIGHT AND TERRY C. GLEASON

UNIVERSITY OF MICHIGAN

An algorithm is presented for constructing from the adjacency matrix of a digraph the matrix of its simple n-sequences. In this matrix, the i, j entry, $i \neq j$, gives the number of paths of length n from a point v_i to a point v_j; the diagonal entry i, i gives the number of cycles of length n containing v_i. The method is then generalized to networks—that is, digraphs in which some value is assigned to each line. With this generalized algorithm it is possible, for a variety of value systems, to calculate the values of the paths and cycles of length n in a network and to construct its value matrix of simple n-sequences. The procedures for obtaining the two algorithms make use of properties of a line digraph—that is, a derived digraph whose points and lines represent the lines and adjacency of lines of the given digraph.

In research on such topics as cognition, learning, verbal behavior, communication, sociometry, and social interaction, empirical structures are often represented by digraphs (directed graphs) in which each point corresponds to an empirical entity and each directed line corresponds to an empirical relationship. A problem frequently encountered in working with digraphs is to find the number of ways one can go from one point to another, using a given number of lines, without passing through any point more than once. Thus, for example, in the context of communication research, one may want to know how many ways a message can go from one person to another through a network in exactly n steps while satisfying the requirement that no person hear the message more than once. Stated in the terminology of digraph theory, the problem is to find the number of paths of length n from one point to another. In solving this problem, it is necessary to deal also with (directed) cycles. We say that a point lies on a cycle of length n if it is possible to leave a point and then return to it in exactly n steps without passing through any other point more than once.

It is well known that if a digraph D contains no cycles, the number of paths of length n from a point v_i to a point v_j is given by the i, j entry of A^n, where A is the adjacency matrix of D. But if D has any cycles, the number of paths of length n cannot be ascertained directly from A^n. There have been several attempts to overcome this limitation. Luce and Perry [4] showed

*The research reported here was supported by Grant NSF-G-17771 from the National Science Foundation. We wish to thank Professor Frank Harary for suggesting certain ways of improving an earlier draft of this paper.

how to find the paths of length 3 in any digraph. Ross and Harary [6] extended the solution to lengths 4, 5, and 6 and presented an algorithm by which any formula for longer paths could eventually be obtained. However, the formula for length 6 is quite formidable and there seems to be little likelihood that a general solution is practicable by their method. Parthasarathy [5] has offered a solution in principle but its usefulness is limited by the great amount of calculation required.

In this paper we again raise the problem and present a method for finding both the number of paths and cycles of any given length through a series of reasonably simple matrix operations. The crux of the solution lies in exploiting the properties of line digraphs as developed by Harary and Norman [1]. We begin with a discussion of line digraphs and their relevance to the problem, then present a matrix method which capitalizes on the properties of line digraphs, and conclude by generalizing the method to networks in which each line has an assigned value. (For a systematic treatment of digraph theory, see Harary, Norman, and Cartwright [2].)

Digraphs and Line Digraphs

A *directed graph* (or *digraph*) is a non-empty set V of points and a prescribed subset of the set of all ordered pairs of the members of V. Each of these ordered pairs (v_i , v_j) is called a *line*, which we denote by v_iv_j or by x_{ij}. The first member of the pair is called the *first point* of the line and the second member is called the *second point*. For any digraph we require that there are no lines whose first and second points are the same and every pair of points v_i and v_j has at most one line v_iv_j.

A (point-line) *sequence* in a digraph 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. If L is a sequence $v_i , x_{ij} , v_j , x_{jk} , v_k , \cdots , v_m , x_{mn} , v_n$, we may denote L simply by indicating the order of occurrence of its points: $L = v_i , v_j , v_k , \cdots , v_m , v_n = L_{ijk\cdots mn}$. A sequence is a *path* if all of its points are distinct. A *cycle* consists of a path from a point u to a point v together with the line vu. We say that a sequence is *simple* if it is either a path or a cycle; all other sequences are *redundant*. The *length* of a sequence is the number of occurrences of lines in it. An *n-sequence* is a sequence of length n. We take a point to be a 0-path and a line to be a 1-path.

The *line digraph* $\mathcal{L}(D)$ of a digraph D is a digraph whose points correspond to the lines of D and whose lines are given by the rule: If x_{ij} and x_{mn} are lines of D, then a line is drawn in $\mathcal{L}(D)$ from (the point corresponding to) x_{ij} to x_{mn} if in D the second point of x_{ij} is the same as the first point of x_{mn} (that is, $v_j = v_m$). Since $\mathcal{L}(D)$ is itself a digraph, we may also form its line digraph $\mathcal{L}(\mathcal{L}(D))$. We let $\mathcal{L}(\mathcal{L}(D)) = \mathcal{L}^2(D)$, and in general $\mathcal{L}(\mathcal{L}^{n-1}(D)) = \mathcal{L}^n(D)$.

Let L be a k-sequence of D. From the definition of $\mathcal{L}(D)$ we see that L yields a $(k - 1)$-sequence in $\mathcal{L}(D)$, a $(k - 2)$-sequence in $\mathcal{L}^2(D)$, and in general, a $(k - n)$-sequence in $\mathcal{L}^n(D)$, for $n \leq k$. In particular, each k-sequence of D yields a line in $\mathcal{L}^{k-1}(D)$ and a point in $\mathcal{L}^k(D)$. These observations are illustrated by the digraphs shown in Fig. 1. The 2-path L_{543} of D yields in $\mathcal{L}(D)$ the line from x_{54} to x_{43} and in $\mathcal{L}^2(D)$ the point y_{543} . The 3-cycle L_{1521} of D yields in $\mathcal{L}(D)$ the 2-path—x_{15} , x_{52} , x_{21}—and in $\mathcal{L}^2(D)$ the line from y_{152} to y_{521} . And the redundant 3-sequence L_{5152} of D yields in $\mathcal{L}(D)$ the 2-path—x_{51} , x_{15} , x_{52}—and in $\mathcal{L}^2(D)$ the line from y_{515} to y_{152} .

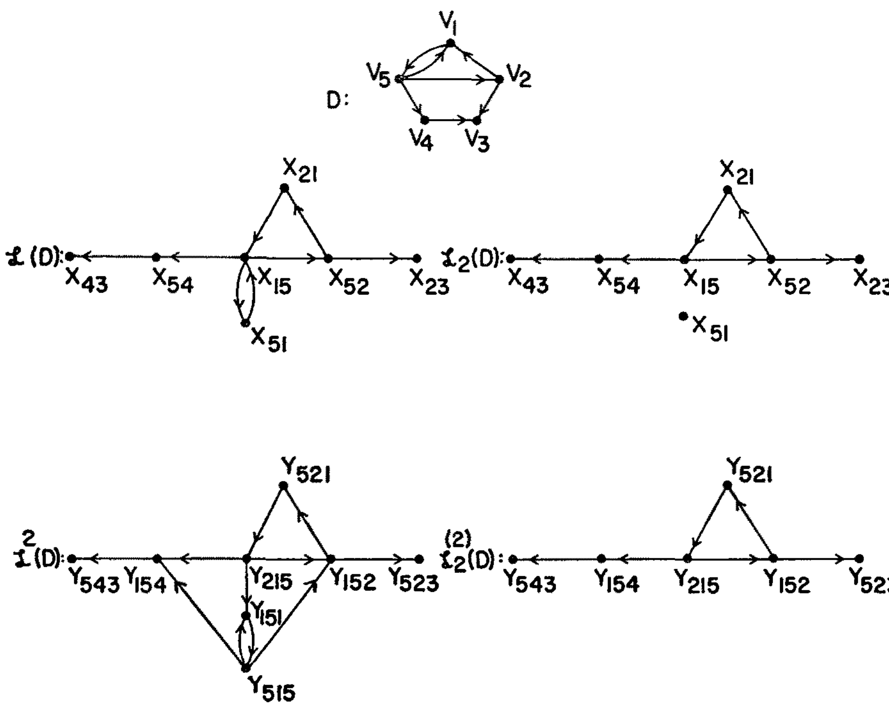

FIGURE 1

Since there is a one-to-one correspondence between k-sequences of D and lines in $\mathcal{L}^{k-1}(D)$, it is possible to ascertain the number of k-sequences from $\mathcal{L}^{k-1}(D)$. It is important to note that if D has no cycles, then every k-sequence of D is a k-path. In this case, the k-paths of D correspond uniquely to the lines of $\mathcal{L}^{k-1}(D)$. But if D has any cycles of length less than k, some of the lines in $\mathcal{L}^{k-1}(D)$ will correspond to redundant sequences. Since we are here interested only in simple sequences, we want a method for obtaining a modified digraph, analogous to $\mathcal{L}^{k-1}(D)$, in which each line corresponds to a simple k-sequence of D, and we want to be able to distinguish between paths and cycles. We now describe such a method, leaving until later a detailed proof of its general applicability.

It is evident that every 2-sequence of a digraph is either a 2-path or a 2-cycle. Let Z be a 2-cycle of D, as for example, the one in Figure 1 containing v_1 and v_5. Now there is a 2-sequence beginning and ending at each of these points, each of which is a 2-cycle said to be *rooted* at the indicated point. The one, L_{515}, is rooted at v_5 and yields in $\mathcal{L}(D)$ the line from x_{51} to x_{15}. The other, L_{151}, is rooted at v_1 and yields the line from x_{15} to x_{51}. Note that these two lines form a 2-cycle in $\mathcal{L}(D)$. On the other hand, a 2-path in D yields one line in $\mathcal{L}(D)$ that does not lie on a 2-cycle. We see, then, that each line in $\mathcal{L}(D)$ corresponds to a simple 2-sequence of D; it represents a 2-cycle of D if it lies on a 2-cycle, and it represents a 2-path otherwise.

We want next to construct from $\mathcal{L}(D)$ a modified digraph, analogous to $\mathcal{L}^2(D)$, whose lines correspond only to simple 3-sequences of D. Since a 3-sequence is simple if and only if it does not contain a 2-cycle, we may remove from $\mathcal{L}(D)$ all lines representing 2-cycles of D, namely, those lying on a 2-cycle (without, however, removing the first and second points of these lines). We denote the resulting digraph by $\mathcal{L}_2(D)$. Clearly, every line in $\mathcal{L}_2(D)$ uniquely corresponds to a 2-path. Let us now form the line digraph of $\mathcal{L}_2(D)$ and denote it by $\mathcal{L}_2^{(2)}(D)$. Each line in $\mathcal{L}_2^{(2)}(D)$ corresponds to two lines in $\mathcal{L}_2(D)$ that have a point in common, which occurs if and only if the lines in $\mathcal{L}_2(D)$ represent two 2-paths of D having a line in common. Two such paths, taken together, form a 3-path if and only if the first point of one is different from the last point of the other. Otherwise, they form a 3-cycle. Thus, every line in $\mathcal{L}_2^{(2)}(D)$ represents a simple 3-sequence of D.

We saw above that rooted 2-cycles of D correspond to lines that lie on a 2-cycle in $\mathcal{L}(D)$. We now show that, in a similar way, each rooted 3-cycle of D yields a line in $\mathcal{L}_2^{(2)}(D)$ that lies on a 3-cycle. Let Z be a 3-cycle of D containing the points v_i, v_j, and v_k. Clearly, there are three 2-paths of D of the form: L_{ijk}, L_{jki}, and L_{kij}. Now, L_{ijk} and L_{jki}, taken together, form a 3-cycle rooted at v_i. They have the line x_{jk} in common and yield in $\mathcal{L}_2^{(2)}(D)$ a line from y_{ijk} to y_{jki}. Similarly, L_{jki} and L_{kij} form a 3-cycle rooted at v_j and yield a line from y_{jki} to y_{kij}. Finally, L_{kij} and L_{ijk} form a 3-cycle rooted at v_k and yield a line from y_{kij} to y_{ijk}. It is clear, then, that each 3-cycle of D yields three lines in $\mathcal{L}_2^{(2)}(D)$ which form a 3-cycle and each line of this 3-cycle represents a rooted 3-cycle of D. Thus, the lines in $\mathcal{L}_2^{(2)}(D)$ provide the desired information concerning the simple 3-sequences of D; each line lying on a 3-cycle represents a rooted 3-cycle of D and every other line corresponds to a 3-path.

These observations are illustrated in Fig. 1. Clearly, there is a one-to-one correspondence between the 2-paths of D and the lines in $\mathcal{L}_2(D)$. There is, furthermore, a one-to-one correspondence between the simple 3-sequences of D and the lines in $\mathcal{L}_2^{(2)}(D)$. Finally, we note that in $\mathcal{L}_2^{(2)}(D)$ each line not lying on a 3-cycle corresponds to a 3-path of D whereas each line lying on a 3-cycle represents a rooted 3-cycle.

The procedure just described can easily be generalized. We take as the inductive hypothesis that there is a one-to-one correspondence between the simple n-sequences of D and the lines of $\mathcal{L}_{n-1}^{(n-1)}(D)$. Let $\mathcal{L}_n^{(n-1)}(D)$ be the subgraph of $\mathcal{L}_{n-1}^{(n-1)}(D)$ obtained by removing the lines of all its n-cycles. Then each line in $\mathcal{L}_n^{(n-1)}(D)$ corresponds to an n-path of D. Let

$$\mathcal{L}_n^{(n)}(D) = \mathcal{L}(\mathcal{L}_n^{(n-1)}(D)).$$

Now each line in $\mathcal{L}_n^{(n)}(D)$ corresponds to two n-paths of D having $n - 1$ lines in common. Clearly, these two paths together form a simple $(n + 1)$-sequence of D. If the first point of one of these paths is the same as the last point of the other, they combine to form an $(n + 1)$-cycle Z; and Z contains $n + 1$ rooted $(n + 1)$-cycles, each of which yields a line lying on an $(n + 1)$-cycle in $\mathcal{L}_n^{(n)}(D)$. Any line in $\mathcal{L}_n^{(n)}(D)$ not lying on an $(n + 1)$-cycle must, therefore, correspond to an $(n + 1)$-path of D.

We conclude, then, that the lines of $\mathcal{L}_{k-1}^{(k-1)}(D)$ provide the desired information concerning simple k-sequences of D. That is, each line lying on a k-cycle uniquely corresponds to a rooted k-cycle of D, and every other line uniquely corresponds to a k-path of D. In the remainder of this paper we show how this information may be obtained by matrix methods.

Matrix Operations

Let D be a digraph whose points are labeled $v_1, v_2, \cdots, v_p$. The *adjacency matrix* of D, $A = A(D)$, is a $p \times p$ matrix whose entries are $a_{ij} = 1$ if there is a line v_iv_j in D and $a_{ij} = 0$ otherwise. The *matrix of simple n-sequences of D*, $S_n = S_n(D)$, is also a $p \times p$ matrix whose entry $s_{ij}^{(n)}$ is the number of distinct n-paths from v_i to v_j when $i \neq j$, and whose entry $s_{ii}^{(n)}$ is the number of n-cycles rooted at v_i. Clearly, $S_1 = A$. Given A, our problem is to find a series of matrix operations which will result in S_n, $2 \leq n \leq p$.

In order to solve this problem we introduce the notion of the adjacency of two sequences of a digraph. If L is a sequence of D, then the first point of L will be denoted $\alpha(L)$, and the last point will be denoted $\omega(L)$. Let L_i and L_j be two sequences of D. We say that L_i is *n-adjacent to* L_j if there is an n-path L_e which is a subpath of both L_i and L_j and which satisfies the conditions: $\omega(L_e) = \omega(L_i)$ and $\alpha(L_e) = \alpha(L_j)$. Now if L_i is a simple n-sequence, then clearly $L_i = L_e$ and L_i is a subpath of L_j containing its first n lines. Similarly, if L_j is a simple n-sequence, then $L_j = L_e$ and L_j is a subpath of L_i consisting of its last n lines. We saw above that there is a line from L_i to L_j in $\mathcal{L}_n^{(n)}(D)$ if and only if L_i and L_j are n-paths of D having $n - 1$ lines in common. In the present terminology, this means that each line from L_i to L_j in $\mathcal{L}_n^{(n)}(D)$ corresponds to two n-paths L_i and L_j of D such that

L_i is $(n-1)$-adjacent to L_j. In the following discussion, we are concerned with the adjacency of the simple sequences of a digraph D.

Let $\mathfrak{A}$ be a square matrix whose rows and columns correspond to the simple n-sequences in D, for all $n = 0, 1, \cdots, p$. We assume an ordering θ of all simple sequences, and hence of the rows and columns of $\mathfrak{A}$, which is subject only to the restriction that every n-sequence precedes every $(n+1)$-sequence. By θ_n we shall mean θ restricted to the simple n-sequences. We denote by A_{mn} the submatrix of $\mathfrak{A}$ whose rows correspond to simple m-sequences and whose columns correspond to simple n-sequences. Clearly, $\mathfrak{A}$ can be expressed in terms of its submatrices as follows:

$$\mathfrak{A} = \begin{bmatrix} A_{00} & A_{01} & \cdots & A_{0n} & \cdots & A_{0p} \\ A_{10} & A_{11} & \cdots & A_{1n} & \cdots & A_{1p} \\ \vdots & \vdots & & \vdots & & \vdots \\ A_{m0} & A_{m1} & \cdots & A_{mn} & \cdots & A_{mp} \\ \vdots & \vdots & & \vdots & & \vdots \\ A_{p0} & A_{p1} & \cdots & A_{pn} & \cdots & A_{pp} \end{bmatrix}.$$

The entries $a_{ij}^{(mn)}$ of A_{mn} are given by the rules

$$a_{ij}^{(mn)} = 1 \text{ if } \begin{cases} m < n \text{ and } L_i \text{ is } m\text{-adjacent to } L_j, \\ m > n \text{ and } L_i \text{ is } n\text{-adjacent to } L_j, \\ m = n,\ n > 0, \text{ and } L_i \text{ and } L_j \text{ are } n\text{-paths such that} \\ \qquad L_i \text{ is } (n-1)\text{-adjacent to } L_j, \\ m = n = 0 \text{ and the line } v_iv_j \text{ is in } D; \end{cases}$$

otherwise, $a_{ij}^{(mn)} = 0$.

Note that for a given ordering θ, the submatrix A_{00} is the adjacency matrix of D, and the entire matrix $\mathfrak{A}$ is well defined. We now show how the matrix S_n can be obtained from certain submatrices of $\mathfrak{A}$, leaving until later the actual construction of $\mathfrak{A}$ itself.

Consider the submatrix A_{0n} for $n > 0$. Its ith row corresponds to the point v_i in D, its kth column corresponds to the simple n-sequence L_k, and the entry $a_{ik}^{(0n)} = 1$ if and only if $v_i = \alpha(L_k)$. Likewise in the submatrix A_{n0}, $a_{kj}^{(n0)} = 1$ if and only if $\omega(L_k) = v_j$. Thus, in the product $A_{0n}A_{n0}$ the term

$$a_{ik}^{(0n)}a_{kj}^{(n0)} = 1$$

if and only if L_k is a simple n-sequence from v_i to v_j. If $v_i = v_j$, L_k is an n-cycle rooted at v_i; otherwise L_k is an n-path. Since the matrix $\mathfrak{A}$ is defined so that all the simple n-sequences in D are represented in the columns of A_{0n}

and the rows of A_{n0}, we see immediately that

$$\sum_k a_{ik}^{(0n)} a_{kj}^{(n0)} = m$$

if and only if there are m simple n-sequences from v_i to v_j (i.e., $s_{ij}^{(n)} = m$). This essentially proves our first theorem.

THEOREM 1. *For all $n > 0$, the matrix of simple n-sequences is given by $S_n = A_{0n}A_{n0}$.*

$\mathfrak{A}$ =

v_1	0 0 0 0 1	1 0 0 0 0 0 0	1 1 1 0 0 0 0 0	1 1 1 0 0 0	0
v_2	1 0 1 0 0	0 1 1 0 0 0 0	0 0 0 1 0 0 0 0	0 0 0 1 1 0	1
v_3	A_{00}: 0 0 0 0 0	A_{01}: 0 0 0 0 0 0 0	A_{02}: 0 0 0 0 0 0 0 0	A_{03}: 0 0 0 0 0 0	A_{04}: 0
v_4	0 0 1 0 0	0 0 0 1 0 0 0	0 0 0 0 0 0 0 0	0 0 0 0 0 0	0
v_5	1 1 0 1 0	0 0 0 0 1 1 1	0 0 0 0 1 1 1 1	0 0 0 0 0 1	0
L_{15}	0 0 0 0 1	0 0 0 0 1 1 1	1 1 1 0 0 0 0 0	1 1 1 0 0 0	0
L_{21}	1 0 0 0 0	1 0 0 0 0 0 0	0 0 0 1 0 0 0 0	0 0 0 1 1 0	1
L_{23}	0 0 1 0 0	0 0 0 0 0 0 0	0 0 0 0 0 0 0 0	0 0 0 0 0 0	0
L_{43}	A_{10}: 0 0 1 0 0	A_{11}: 0 0 0 0 0 0 0	A_{12}: 0 0 0 0 0 0 0 0	A_{13}: 0 0 0 0 0 0	A_{14}: 0
L_{51}	1 0 0 0 0	1 0 0 0 0 0 0	0 0 0 0 1 0 0 0	0 0 0 0 0 0	0
L_{52}	0 1 0 0 0	0 1 1 0 0 0 0	0 0 0 0 0 1 1 0	0 0 0 0 0 1	0
L_{54}	0 0 0 1 0	0 0 0 1 0 0 0	0 0 0 0 0 0 0 1	0 0 0 0 0 0	0
L_{151}	1 0 0 0 0	0 0 0 0 1 0 0	0 0 0 0 0 0 0 0	0 0 0 0 0 0	0
L_{152}	0 1 0 0 0	0 0 0 0 0 1 0	0 0 0 0 0 1 1 0	1 1 0 0 0 0	0
L_{154}	0 0 0 1 0	0 0 0 0 0 0 1	0 0 0 0 0 0 0 1	0 0 1 0 0 0	0
L_{215}	0 0 0 0 1	1 0 0 0 0 0 0	0 1 1 0 0 0 0 0	0 0 0 1 1 0	1
L_{515}	A_{20}: 0 0 0 0 1	A_{21}: 1 0 0 0 0 0 0	A_{22}: 0 0 0 0 0 0 0 0	A_{23}: 0 0 0 0 0 0	A_{24}: 0
L_{521}	1 0 0 0 0	0 1 0 0 0 0 0	0 0 0 1 0 0 0 0	0 0 0 0 0 1	0
L_{523}	0 0 1 0 0	0 0 1 0 0 0 0	0 0 0 0 0 0 0 0	0 0 0 0 0 0	0
L_{543}	0 0 1 0 0	0 0 0 1 0 0 0	0 0 0 0 0 0 0 0	0 0 0 0 0 0	0
L_{1521}	1 0 0 0 0	0 1 0 0 0 0 0	0 0 0 0 0 1 0 0	0 0 0 0 0 0	0
L_{1523}	0 0 1 0 0	0 0 1 0 0 0 0	0 0 0 0 0 0 1 0	0 0 0 0 0 0	0
L_{1543}	0 0 1 0 0	0 0 0 1 0 0 0	0 0 0 0 0 0 0 1	0 0 0 0 0 0	0
L_{2152}	A_{30}: 0 1 0 0 0	A_{31}: 0 0 0 0 0 1 0	A_{32}: 0 1 0 0 0 0 0 0	A_{33}: 0 0 0 0 0 0	A_{34}: 0
L_{2154}	0 0 0 1 0	0 0 0 0 0 0 1	0 0 1 0 0 0 0 0	0 0 1 0 0 0	1
L_{5215}	0 0 0 0 1	1 0 0 0 0 0 0	0 0 0 1 0 0 0 0	0 0 0 0 0 0	0
L_{21543}	A_{40}: 0 0 1 0 0	A_{41}: 0 0 0 1 0 0 0	A_{42}: 0 0 0 0 0 0 0 1	A_{43}: 0 0 1 0 0 0	A_{44}: 0

FIGURE 2

Fig. 2 gives the matrix $\mathfrak{A}$ for the digraph D of Fig. 1. For convenience, each row is labeled so as to identify the corresponding simple sequence and, of course, the kth row and the kth column correspond to the same simple sequence. Thus, for example, the entry $a_{24}^{(02)} = 1$ indicates that v_2 is the first point of L_{215} and the entry $a_{45}^{(20)} = 1$ means that v_5 is the last point of L_{215}.

Upon multiplying the appropriate submatrices of Fig. 2, we obtain the following matrices which give the number of n-paths and n-cycles, for each n, in the digraph of Fig. 1.

$$A_{02}A_{20} = S_2 = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 \end{bmatrix},$$

$$A_{03}A_{30} = S_3 = \begin{bmatrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$A_{04}A_{40} = S_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The only nonzero entries on the main diagonal of S_2 are $s_{11}^{(2)} = s_{55}^{(2)} = 1$. This tells us that D has only one 2-cycle and that it contains v_1 and v_5. The entry $s_{53}^{(2)} = 2$ means that there are two 2-paths from v_5 to v_3, and inspection of D shows these to be L_{543} and L_{523}. In S_3 the entries on the main diagonal indicate that D has one 3-cycle containing v_1, v_2, and v_5. The remaining nonzero entries give the number of 3-paths from v_i to v_j. Finally, S_4 reveals that D has only one 4-path and that it goes from v_2 to v_3.

We have seen that the matrix S_n can be obtained from the submatrices A_{0n} and A_{n0} of $\mathfrak{A}$. Before describing how the entire matrix $\mathfrak{A}$ can be constructed from the adjacency matrix A_{00}, we need to establish a number of structural relations among the submatrices of $\mathfrak{A}$.

We begin by considering A_{nn} and $A_{n,n+1}$. The next theorem, concerning the adjacency of simple sequences in a digraph, shows that the unit entries of these two submatrices are related in a particular way.

THEOREM 2. *In any digraph, the number of ordered pairs of n-paths such that the first is $(n - 1)$-adjacent to the second is equal to the number of ordered pairs consisting of a simple n-sequence followed by a simple $(n + 1)$-sequence such that the first is n-adjacent to the second.*

PROOF. Let L_i and L_j be any two n-paths such that L_i is $(n - 1)$-adjacent to L_j. Clearly, there is an $(n + 1)$-sequence, L_g, which begins at $\alpha(L_i)$ and contains the line from $\alpha(L_i)$ to $\alpha(L_j)$ and the path L_j. Since all points of L_i and of L_j are distinct, L_g is simple. Moreover, there is just one $(n + 1)$-sequence that contains both L_i and L_j. Now since $\alpha(L_i) = \alpha(L_g)$ and L_g contains L_i, it follows that L_i is n-adjacent to L_g. Thus, for each ordered pair, L_i and L_j, there is a unique ordered pair, L_i and L_g, such that L_i is n-adjacent to L_g.

Now assume that there exists an n-path L_i which is n-adjacent to the simple $(n + 1)$-sequence L_g. Let L_h be the first line of L_g and let L_j be the subsequence of L_g from $\omega(L_h)$ to $\omega(L_g)$. Since the points of L_g, with the possible exception of $\alpha(L_g)$ and $\omega(L_g)$, are distinct, L_j is an n-path. Moreover, L_i and L_j have a common $(n - 1)$-path L_e. Since $\alpha(L_e) = \alpha(L_j)$ and $\omega(L_e) = \omega(L_i)$, it follows that L_i is $(n - 1)$-adjacent to L_j. There can be only one such pair, L_i and L_j, since any simple $(n + 1)$-sequence contains exactly one pair of n-paths. □

Recall that (by definition) the entry $a_{ij}^{(nn)} = 1$ indicates that the n-path L_i is $(n - 1)$-adjacent to the n-path L_j, and the entry $a_{ij}^{(n,n+1)} = 1$ indicates that the simple n-sequence L_i is n-adjacent to the simple $(n + 1)$-sequence L_j. The following result thus follows immediately from Theorem 2.

COROLLARY 2a. *The number of unit entries in A_{nn} equals the number of unit entries in $A_{n,n+1}$.*

It is readily apparent that for each simple $(n + 1)$-sequence in D there is exactly one n-path that is n-adjacent to it. In other words, each column in $A_{n,n+1}$ contains exactly one unit entry, and we may conclude that the number of columns in $A_{n,n+1}$ equals the number of unit entries in A_{nn}. We see, then, that there is a one-to-one correspondence between the unit entries in A_{nn} and the simple $(n + 1)$-sequences of D. Hence any ordering, θ_{n+1}, of one of these induces an ordering of the other.

Let us assume now that the $r \times r$ matrix A_{nn} is known, for some n. By this assumption, there is an ordering θ_n of the simple n-sequences of D and hence of the corresponding rows and columns of $\mathfrak{A}$. Let us arbitrarily order the unit entries of A_{nn} by θ_{n+1}. The submatrix $A_{n,n+1}$ can then be obtained from A_{nn} in the following way.

Construction 1. If the kth unit entry, relative to θ_{n+1}, of A_{nn} occurs in its ith row, then $A_{n,n+1}$ is obtained by letting its kth column be the $r \times 1$ vector which has 1 in its ith place and 0's elsewhere.

Clearly this construction satisfies the definition of $A_{n,n+1}$, for its rows and columns are ordered by θ_n and θ_{n+1} respectively and there is a 1 in the i, k entry of the constructed matrix if and only if the n-path L_i is n-adjacent to the simple $(n + 1)$-sequence L_k.

The matrix $A_{n+1,n}$ can be constructed from A_{nn} in a similar way. By an argument analogous to that in the proof of Theorem 2 it can be shown that the number of unit entries in A_{nn} is the same as the number of unit entries in $A_{n+1,n}$ and there is exactly one unit entry in each row of $A_{n+1,n}$. We may, then, employ the same ordering θ_{n+1} as above and construct $A_{n+1,n}$ from A_{nn}.

Construction 2. If the kth unit entry, relative to θ_{n+1}, of A_{nn} occurs in its jth column, then $A_{n+1,n}$ is obtained by letting its kth row be the $1 \times r$ vector which has 1 in its jth place and 0's elsewhere.

By this construction, there is a 1 in the k, j entry of the obtained matrix if and only if the simple $(n + 1)$-sequence L_k is n-adjacent to the n-path L_j. Thus, the definition of $A_{n+1,n}$ is satisfied.

(The submatrices under consideration contain information relevant to our earlier discussion of line digraphs. The submatrix A_{nn} contains the adjacency matrix of $\mathcal{L}_n^{(n)}(D)$, as can be seen by comparing Figs. 1 and 2. And if the transpose of a matrix M is denoted M', then $A_{10} + A'_{01}$ is obtained from the incidence matrix of D by replacing every nonzero entry by $+1$. Likewise, $A_{n+1,n} + A'_{n,n+1}$ contains the incidence matrix of $\mathcal{L}_n^{(n)}(D)$ as a submatrix.)

These two constructions may be illustrated by Fig. 2. The unit entry $a_{26}^{(22)}$ indicates that L_{152} is 1-adjacent to L_{521}. We let θ_2 be the ordering of the unit entries of A_{22} obtained in the following way. Start with the first row and read from left to right, noting the order of occurrence of unit entries; go to the second row and read from left to right, and continue to note the order of occurrence of unit entries; continue this procedure through the last row. We see that $a_{26}^{(22)}$ is the first unit entry. Thus, the first column of A_{23} has a 1 in its second place and zeros elsewhere. Similarly, the first row of A_{32} has a 1 in its sixth place and zeros elsewhere. The unit entry $a_{21}^{(23)}$ means that L_{152} is 2-adjacent to a simple 3-sequence, namely, L_{1521}. And the entry $a_{16}^{(32)}$ means that L_{1521} is 2-adjacent to a 2-path, namely, L_{521}.

We next consider how the matrix A_{nn} is related to $A_{n,n-1}$ and $A_{n-1,n}$. By definition, the rows and columns of A_{nn} correspond to the simple n-sequences in D, and $a_{ij}^{(nn)} = 1$ if and only if L_i and L_j are n-paths and L_i is $(n - 1)$-adjacent to L_j. In the product $A_{n,n-1}A_{n-1,n}$, the rows and columns also correspond to the simple n-sequences in D, and each i, j entry is given by

$$\sum_e a_{ie}^{(n,n-1)} a_{ej}^{(n-1,n)}.$$

Now, $a_{ie}^{(n,n-1)} a_{ej}^{(n-1,n)} = 1$ if and only if L_e is an $(n - 1)$-path such that L_i is $(n - 1)$-adjacent to L_e and L_e is $(n - 1)$-adjacent to L_j, or equivalently L_i is $(n - 1)$-adjacent to L_j. Since for any pair, L_i and L_j, there can be only one L_e which satisfies the above conditions, every entry in $A_{n,n-1}A_{n-1,n}$ is either 0 or 1. Thus the only difference between $A_{n,n-1}A_{n-1,n}$ and A_{nn} is that unit entries in the former represent the $(n - 1)$-adjacency of two simple n-sequences whereas in the latter they represent the $(n - 1)$-adjacency of two n-paths. Hence if in $A_{n,n-1}$ and $A_{n-1,n}$ we change to 0 all the unit entries involving n-cycles and denote the resulting matrices by $B_{n,n-1}$ and $B_{n-1,n}$, then A_{nn} is obtained by forming the product $B_{n,n-1}B_{n-1,n}$. These observations constitute a proof of the next theorem. (Since the unit entries in $B_{n,n-1}$ and $B_{n-1,n}$ represent only n-paths, use of these matrices has the effect of removing n-cycles from $\mathcal{L}_{n-1}^{(n-1)}(D)$.)

THEOREM 3. *For all $n \geq 1$, the matrix A_{nn} is the product*

$$A_{nn} = B_{n,n-1}B_{n-1,n}.$$

The next theorem shows that $B_{n,n-1}$ and $B_{n-1,n}$ can be obtained by matrix operations which identify the unit entries involving n-cycles in $A_{n,n-1}$ and $A_{n-1,n}$. We rely on two matrix operations not previously used: (a) the transpose of M, denoted M', and (b) the element-wise product $M \times N$, where both M and N have r rows and c columns and the i, j entry of $M \times N$ is given by $m_{ij} \cdot n_{ij}$.

Since our definition of a digraph precludes 1-cycles, it is immediate that $A_{10} = B_{10}$ and $A_{01} = B_{01}$.

THEOREM 4. *For all $n \geq 2$, the matrices $B_{n,n-1}$ and $B_{n-1,n}$ are given by the equations*

$$B_{n,n-1} = A_{n,n-1} - [A_{n,n-1} \times (A_{n-1,0}A_{0n})'],$$

and

$$B_{n-1,n} = A_{n-1,n} - [A_{n-1,n} \times (A_{n0}A_{0,n-1})'].$$

PROOF. Consider first the matrix $(A_{n-1,0}A_{0n})'$, whose entries are given by

$$\left(\sum_e a_{je}^{(n-1,0)} a_{ei}^{(0,n)}\right)'.$$

Now, $a_{je}^{(n-1,0)} = 1$ if and only if the simple $(n - 1)$-sequence L_j is 0-adjacent to the 0-sequence L_e, that is $\omega(L_j) = L_e$. And $a_{ei}^{(0,n)} = 1$ if and only if L_e is 0-adjacent to the simple n-sequence L_i, that is, $\alpha(L_i) = L_e$. Clearly, there can be at most one point L_e satisfying both these conditions. Hence, every entry in $(A_{n-1,0}A_{0n})'$ is either 0 or 1, and each unit entry, i, j, indicates that $\alpha(L_i) = \omega(L_j)$. By definition, $a_{ij}^{(n,n-1)} = 1$ if and only if L_i is $(n - 1)$-adjacent to L_j, that is, $\omega(L_i) = \omega(L_j)$. Since the transpose of $(A_{n-1,0}A_{0n})$ has the same number of rows and columns as $A_{n-1,n}$, we may form the element-wise product of these matrices to obtain the matrix given by the bracketed term in the first equation of the theorem. Now, the i, j entry of this matrix is 1 if and only if $\alpha(L_i) = \omega(L_j)$ and $\omega(L_i) = \omega(L_j)$ or, equivalently, $\alpha(L_i) = \omega(L_i)$. Since L_i is a simple n-sequence, it follows that L_i is an n-cycle. Thus, each i, j unit entry in this matrix corresponds to an n-cycle L_i that is $(n - 1)$-adjacent to an $(n - 1)$-sequence L_j. Upon subtracting this matrix from $A_{n,n-1}$, we obtain a matrix satisfying the definition of $B_{n,n-1}$.

A similar argument establishes the second equation of the theorem. □

Up to this point we have considered how the submatrices A_{nn}, which lie on the main diagonal of $\mathcal{A}$, are related to submatrices immediately adjacent to the main diagonal. We now show that any submatrix A_{mn}, $m \neq n$, is related in a particular way to certain other submatrices of $\mathcal{A}$. The next two lemmas concerning the adjacency of sequences in a digraph provide a basis for establishing these relationships.

LEMMA 1. *Let L_0, L_1, $\cdots$, L_r be a series of sequences such that L_e is an $(m + e)$-sequence, for all $e = 0, 1, \cdots, r$. If L_e is $(m + e)$-adjacent to L_{e+1}, for all $e = 0, 1, \cdots, r - 1$, then L_0 is m-adjacent to L_r.*

PROOF. If L_0 is m-adjacent to L_1 and L_1 is $(m + 1)$-adjacent to L_2, then L_1 must be an $(m + 1)$-path containing L_0. But since L_1 is contained in L_2 and $\alpha(L_0) = \alpha(L_1) = \alpha(L_2)$, we see that L_0 is m-adjacent to L_2. Now, if L_0 is m-adjacent to L_{r-1} and L_{r-1} is $(m + r - 1)$-adjacent to L_r, then L_{r-1} must be an $(m + r - 1)$-path contained in L. Since

$$\alpha(L_0) = \alpha(L_{r-1}) = \alpha(L_r),$$

it follows that L_0 is m-adjacent to L_r. □

The next lemma is in the nature of a converse to Lemma 1.

LEMMA 2. *Let L_0 be an m-path and L_r a simple n-sequence. If L_0 is m-adjacent to L_r, then there exists a unique series of sequences L_0, L_1, $\cdots$, L_r such that L_e is $(m + e)$-adjacent to L_{e+1}, for all $e = 0, 1, \cdots, r - 1$.*

PROOF. For any $e = 0, 1, \cdots, r$ let L_{e+1} be the $(m + e)$-subpath of L_r for which $\alpha(L_e) = \alpha(L_r)$. Since L_e is contained in L_{e+1} and $\alpha(L_e) = \alpha(L_{e+1})$, L_e is $(m + e)$-adjacent to L_{e+1}. Thus the paths L_e form a series L_0, L_1, $\cdots$, L_r. That this series is unique follows from the fact that L_e is the only $(m + e)$-subpath of L_r for which $\alpha(L_e) = \alpha(L_r)$. □

THEOREM 5. *For any $m < n$,*

$$(1) \qquad A_{mn} = A_{m,m+1}A_{m+1,m+2} \cdots A_{n+1,n},$$

$$(2) \qquad A_{nm} = A_{n,n-1}A_{n-1,n-2} \cdots A_{m+1,m}.$$

PROOF. The entries of the matrix $A_{m,m+1}A_{m+1,m+2} \cdots A_{n-1,n}$ are composed of products of the form $a_{if}^{(m,m+1)}a_{fg}^{(m+1,m+2)} \cdots a_{hj}^{(n-1,n)}$. Clearly this product must be 0 or 1. If it equals 1, then there is an n-sequence L_i in D such that L_i is m-adjacent to the $(m + 1)$-sequence L_f. Likewise, L_f is $(m + 1)$-adjacent to the $(m + 2)$-sequence in L_g. This series continues until we come to the $(n - 1)$-sequence L_h which is $(m + n - 1)$-adjacent to the simple n-sequence L_j. By Lemma 1, L_i is m-adjacent to L_j. By Lemma 2, the series described in the above product is unique. Thus the i, j entry of $A_{m,m+1}A_{m+1,m+2} \cdots A_{n+1,n}$ is either 0 or 1, and if it is 1, then $a_{ij}^{(mn)} = 1$.

If $a_{ij}^{(mn)} = 1$, then L_i is m-adjacent to L_j. By Lemma 2 we may construct a series $L_i, L_1, L_2, \cdots, L_e, \cdots, L_j$ such that L_e is $(m + e)$-adjacent to L_{e+1}. Thus $a_{i1}^{(m,m+1)}a_{12}^{(m+1,m+2)} \cdots a_{rj}^{(n-1,n)} = 1$, and since this series is unique the i, j entry of $A_{m,m+1}A_{m+1,m+2} \cdots A_{n-1,n}$ is 1.

This establishes the validity of statement (1) of the theorem. A similar argument holds for statement (2). □

It will be recalled that Theorem 1 shows that the matrix S_n can be obtained from the submatrices A_{0n} and A_{n0}. The following corollary of Theorem 5 therefore provides information useful in finding S_n.

COROLLARY 5a. *For any* $n = 2, 3, \cdots, p$,

$$(1) \qquad A_{0n} = A_{0,n-1}A_{n-1,n},$$

$$(2) \qquad A_{n0} = A_{n,n-1}A_{n-1,0}.$$

Proof. By Theorem 5,

$$A_{0n} = A_{01}A_{12} \cdots A_{n-2,n-1}A_{n-1,n}.$$

But

$$A_{01}A_{12} \cdots A_{n-2,n-1} = A_{0,n-1}.$$

This establishes (1), and a similar argument establishes (2). □

We summarize the above material and the constructions involved by the following algorithm.

Algorithm 1. Let D be a digraph and θ_n be an arbitrary ordering of the simple n-sequences of D. Then the matrix S_n of simple n-sequences of D, for $n > 0$, can be obtained by the following procedure.

1. Order the points of D by θ_0 and construct A_{00}.
2. Order the unit entries of A_{00} by θ_1 and construct A_{01} and A_{10}.
3. Construct A_{11} by the formula

$$A_{11} = A_{10}A_{01}.$$

Steps 4–7 give a recursive procedure for finding A_{nn}, $n = 2, 3, \cdots, p$.

4. Order the unit entries of $A_{n-1,n-1}$ by θ_n and construct $A_{n-1,n}$ and $A_{n,n-1}$.
5. Construct A_{0n} and A_{n0} by the formulas

$$A_{0n} = A_{0,n-1}A_{n-1,n},$$

$$A_{n0} = A_{n,n-1}A_{n-1,0}.$$

6. Construct $B_{n-1,n}$ and $B_{n,n-1}$ by the formulas

$$B_{n-1,n} = A_{n-1,n} - [A_{n-1,n} \times (A_{n0}A_{0,n-1})'],$$

$$B_{n,n-1} = A_{n,n-1} - [A_{n,n-1} \times (A_{n-1,0}A_{0n})'].$$

7. Construct A_{nn} by the formula

$$A_{nn} = B_{n,n-1}B_{n-1,n}.$$

8. If $A_{nn} = [0]$ or $n = p$, terminate the procedure. Otherwise, return to step 4.
9. Construct S_n from the results of step 5 by the formula

$$S_n = A_{0n}A_{n0}$$

Clearly, Step 1 of this algorithm is simply the customary way of constructing the adjacency matrix of a digraph. Steps 2 and 4 are justified by Corollary 2a, together with Constructions 1 and 2. Steps 3 and 7 make use of Theorem 3. Step 5 is justified by Corollary 5a. Step 6 results from Theorem 4. And, finally, Step 9 is given by Theorem 1.

As noted above, the matrix $\mathcal{A}$ of Figure 2 is obtained from the digraph of Fig. 1. In constructing $\mathcal{A}$, we have used a standard procedure for the ordering θ_{n+1} of the unit entries of A_{nn}, for all $n > 0$, as follows. Start with the first row and read from left to right, noting the order of occurrences of unit entries; go to the second row and read from left to right and continue to note the order of occurrence of unit entries; continue similarly through the last row. For completeness, we have presented all submatrices of $\mathcal{A}$, but not all of these are required for finding the matrix S_n of simple n-sequences of D. In fact, the algorithm yields only submatrices of $\mathcal{A}$ lying on the first row (that is, A_{0n}), the first column (A_{n0}), the main diagonal (A_{nn}), and the diagonals adjacent to the main diagonal ($A_{n,n+1}$ and $A_{n+1,n}$).

In using the algorithm to obtain S_n for even moderately large digraphs, Step 6 involves considerable calculation. Our final theorem shows that under certain conditions this step may be eliminated.

THEOREM 6. *If the number of nonzero rows in* A_{0n} *is less than* n, *then*

$$A_{nn} = A_{n,n-1}A_{n-1,n}.$$

PROOF. The entries of A_{0n} indicate which points in D are the first points of the simple n-sequences corresponding to the columns of A_{0n}. If there is a k-cycle in D for $k \geq n$, then there must be at least n points which are the first points of simple n-sequences. Thus if the number of nonzero rows is less than n, there can be no k-cycles in D for all $k \geq n$. Hence, $B_{n,n-1} = A_{n,n-1}$ and $B_{n-1,n} = A_{n-1,n}$, and the equation of the theorem follows from Theorem 3. □

Some Related Matrices

Once the matrix S_n of simple n-sequences of a digraph is obtained for each n, it is relatively easy to construct other useful matrices. We now briefly discuss some of these.

The *matrix of simple sequences* of a digraph D, $S(D)$, is a $p \times p$ matrix whose entry s_{ij} is the number of distinct paths (of any length) in D from v_i to v_j when $i \neq j$, and whose entry s_{ii} is the number of cycles containing v_i. Clearly, then, we have

$$\sum_{n=1}^{p-1} S_n = S(D).$$

The *distance from v_i to v_j*, denoted d_{ij}, is the length of a shortest path from v_i to v_j. If there is no path from v_i to v_j, we let $d_{ij} = \infty$. The *distance matrix* of D, denoted $N(D)$, is a $p \times p$ matrix whose entries are the distances d_{ij}. The following statements are readily established.

1. Every diagonal entry d_{ii} of $N(D)$ is 0.
2. $d_{ij} = \infty$ if there is no entry $s_{ij}^{(n)} \neq 0$ in S_n, for $n \leq p - 1$.
3. Otherwise, d_{ij} is the smallest value of n such that $s_{ij}^{(n)} \neq 0$.

To construct the distance matrix $N(D)$ for the digraph of Fig. 1, we begin by entering 0's on the main diagonal. Next, we transfer the unit entries in A_{00} to $N(D)$. Then we enter 2's in all empty locations of $N(D)$ whenever there is a nonzero entry in S_2. The next step is to enter 3's in all empty locations whenever there is a nonzero entry in S_3. Since each nondiagonal entry in S_4 is 0, we complete $N(D)$ by entering ∞ in the remaining empty locations. The resulting matrix is

$$N(D) = \begin{bmatrix} 0 & 2 & 3 & 2 & 1 \\ 1 & 0 & 1 & 3 & 2 \\ \infty & \infty & 0 & \infty & \infty \\ \infty & \infty & 1 & 0 & \infty \\ 1 & 1 & 2 & 1 & 0 \end{bmatrix}.$$

A *detour from v_i to v_j* is a path of maximum length from v_i to v_j. The *detour matrix* of D is a $p \times p$ matrix $E(D) = [e_{ij}]$, where $e_{ij} = \infty$ if there is no path from v_i to v_j and otherwise e_{ij} is the length of a detour from v_i to v_j. To construct $E(D)$ we use the following immediate observations.

1. Every diagonal entry e_{ii} is 0.
2. $e_{ij} = \infty$ if there is no entry $s_{ij}^{(n)} \neq 0$ in S_n, for $n \leq p - 1$.
3. Otherwise, e_{ij} is the maximum integer n such that $s_{ij}^{(n)} \neq 0$.

A Generalization of the Method

An (*irreflexive*) *network N* is a digraph together with an assignment of some value to each of its lines. Such a value, for example, may be a probability, an integer, or a positive or negative sign. Thus, the value of a line may represent the probability of going directly from one state to another, the cost of traveling directly from one place to another, or the signed quality of affection of one person for another. (For a discussion of various interpretations of networks, see Harary, Norman, and Cartwright ([2], ch. 12 and 13).) Given a network with a particular value system for its lines, the question arises as to how to assign values to its sequences.

If, for example, the values assigned to the lines of a network are probabilities, it is customary to define the value of a sequence from v_i to v_j as the product of the values of its lines, which if the values are independent is the probability of going from v_i to v_j via this sequence. And if there are several sequences from v_i to v_j, we add the values of these sequences to obtain the probability of going from v_i to v_j by any of them.

Or, if the values of a network represent costs, the value of a sequence would be obtained by summing the values of its lines to give the cost of traveling along the entire sequence. Now, if there were several sequences from v_i to v_j and one wanted to know the minimum cost of traveling from v_i to v_j, one would find the minimum value of all these sequences.

In this section, we present a generalization of Theorem 1 which, for a variety of value systems, allows one to determine some function of the values of all simple n-sequences in a network.

We begin by presenting a generalized arithmetic, whose specific meaning will depend upon the value system of the network in question. For any network there must be a rule of combination whereby the lines of a sequence are used to form the value of a sequence. We call this *generalized multiplication* and denote it by $\odot$. Thus, if t and t' are the values on the lines of a 2-sequence, $t \odot t'$ would be the value of the sequence itself. Likewise, if there are several sequences between two points, there must be some way of assigning a value which represents the set of all such sequences. Such an assignment is called *generalized addition* and is denoted by $\oplus$. Thus, if t and t' are the values of two sequences from v_i to v_j, their generalized sum is $t \oplus t'$.

We assume that these operations satisfy the following conditions. First, they are associative and commutative. Second, there is an identity element $\ddot{1}$, with respect to generalized multiplication, which follows the rule

$$t \odot \ddot{1} = t. \tag{1}$$

Third, there is an identity element $\ddot{0}$, with respect to generalized addition, which follows the rule

$$t \oplus \ddot{0} = t. \tag{2}$$

Finally, we have the condition

$$t \odot \ddot{0} = \ddot{0}. \tag{3}$$

In using this arithmetic, we shall make use of the generalized product of two matrices $A \odot B$. By $A \odot B$ we shall mean the matrix whose i, j entry is given by the expression

$$(a_{i1} \odot b_{1j}) \oplus (a_{i2} \odot b_{2j}) \oplus \cdots \oplus (a_{ir} \odot b_{rj}) = \ddot{\sum}_{k=1}^{k=r} a_{ik} \odot b_{kj}. \tag{4}$$

The *value matrix* of a network N is the matrix, $M(N) = M$, whose entry m_{ij} is the value of the line from v_i to v_j if such a line is in N and $\ddot{0}$ otherwise. We illustrate the use of our generalized arithmetic for two rather different value systems. Consider first a network whose values are costs. In this case, the entry m_{ij} of the value matrix is the cost of traveling directly from v_i to v_j if this is possible. If this is not possible, $m_{ij} = \infty$. Suppose that we want to find the minimum cost of traveling from v_i to v_j in n steps. Using a procedure introduced by Hasse [3], we use the operation of ordinary addition for generalized multiplication, and we use the procedure of taking the minimum value of a set of values for generalized addition. Thus, in this specification of our generalized arithmetic, the identity element $\ddot{0}$ is ∞, and 1 is 0. Now if we form the product $M \odot M = M^{[2]}$, we get

$$\begin{aligned} m_{ij}^{[2]} &= (m_{i1} \odot m_{1j}) \oplus (m_{i2} \odot m_{2j}) \oplus \cdots \oplus (m_{ip} \odot m_{pj}) \\ &= \min (m_{i1} + m_{1j}\,, \quad m_{i2} + m_{2j}\,, \cdots, m_{ip} + m_{pj}), \end{aligned}$$

which gives the minimum cost of traveling from v_i to v_j in 2 steps, and in general $M^{[n]}$ gives the minimum costs for n steps. As a second example, consider a network whose values are probabilities. Now we use ordinary multiplication for $\odot$, ordinary addition for $\oplus$, and we use 0 and 1 for the identity elements $\ddot{0}$ and $\ddot{1}$, respectively. With these operations $m_{ij}^{[n]}$ gives the probability of going from v_i to v_j by any sequence of length n. It is clear that various specifications of our generalized arithmetic may be employed, depending upon the nature of the value system of a network.

We continue to develop our generalization by introducing three new matrices. First, we have the *value matrix of simple n-sequences* $M(S_n)$ whose i, j entries are the generalized sums of the values of all the simple n-sequences from v_i to v_j in N. Second F_n, $1 \leq n \leq p$, is a square matrix whose diagonal entries $f_{ii}^{(n)}$ are the values of the simple n-sequences L_i, relative to the ordering θ_n, and whose off-diagonal entries $f_{ij}^{(n)}$ are $\ddot{0}$. Finally, G_n, $1 \leq n \leq p$, is a square matrix whose diagonal entries $g_{ii}^{(n)}$ are the values of the last line of the simple n-sequences L_i, relative to the ordering θ_n, and whose entries $g_{ij}^{(n)}$ are $\ddot{0}$.

If we are given the value matrix M for a network N, we can readily obtain the adjacency matrix $A(N)$ by changing every $\ddot{0}$ to 0 and letting all other values equal 1. We can then construct the matrix $\mathfrak{A}$ for the underlying digraph of the network N. In the following discussion, we need to modify $\mathfrak{A}$ as follows: Let $\ddot{A}_{mn}$ be the matrix obtained from A_{mn} by replacing each occurrence of 0 by $\ddot{0}$ and each occurrence of 1 by $\ddot{1}$. The importance of this modification will soon be apparent.

Now suppose that F_n and G_n are known and let

$$H_n = \ddot{A}_{nn} \odot G_n \; . \tag{5}$$

By the definition of $\odot$ we have

$$h_{ij}^{(n)} = \ddot{\sum_k} \ddot{a}_{ik}^{(nn)} \odot g_{kj}^{(n)}. \tag{6}$$

But by the definition of G_n there is only one value of k such that $g_{kj}^{(n)} \neq \ddot{0}$. Hence

$$h_{ij}^{(n)} = \ddot{a}_{ij}^{(nn)} \odot g_{jj}^{(n)}. \tag{7}$$

If $\ddot{a}_{ij}^{(nn)} = \ddot{1}$, then $h_{ij}^{(n)} = g_{jj}^{(n)}$. Consequently the last line of the $(n + 1)$-sequence formed by the n-paths L_i and L_j has the value g_{jj} . Clearly if $\ddot{a}_{ij}^{(nn)} = \ddot{0}$, then L_i is not $(n - 1)$-adjacent to L_j , and the entry $h_{ij}^{(n)} = \ddot{0}$ reflects this fact.

Now let

$$T_{nn} = F_n \odot H_n \; . \tag{8}$$

Again applying the definition of $\odot$ we have

$$t_{ij}^{(nn)} = \ddot{\sum_k} f_{ik}^{(n)} \odot h_{kj}^{(n)}. \tag{9}$$

By definition of F_n there is only one value of k such that $f_{ik}^{(n)} \neq \ddot{0}$. Hence by (9) and (7) we get,

$$\begin{aligned} t_{ij}^{(nn)} &= f_{ii}^{(n)} \odot h_{ij}^{(n)} \\ &= f_{ii}^{(n)} \odot \ddot{a}_{ij}^{(nn)} \odot g_{jj}^{(n)}. \end{aligned} \tag{10}$$

We have seen before that $g_{jj}^{(n)}$ is the value of the last line of the n-path L_j and hence the value of the last line of the simple $(n + 1)$-sequence formed by L_i and L_j . Likewise, $f_{ii}^{(n)}$ is the value of the n-path of L_i and hence the value of the first n lines of the larger sequence. Thus the value of the simple $(n + 1)$-sequence corresponding to $\ddot{a}_{ij}^{(nn)} = \ddot{1}$ is the product

$$\begin{aligned} f_{ii}^{(n)} \odot \ddot{a}_{ij}^{(nn)} \odot g_{jj}^{(n)} &= f_{ii}^{(n)} \odot \ddot{1} \odot g_{jj}^{(n)} \\ &= f_{ii}^{(n)} \odot g_{jj}^{(n)}. \end{aligned} \tag{11}$$

Consequently we see that T_{nn} is a kind of generalization of A_{nn} .

These results give us a means of constructing the matrices F_{n+1} and G_{n+1} from T_{nn} and H_n. Assume that F_n, G_n, and A_{nn} are known. From these matrices we can calculate T_{nn} and H_n according to (5) and (8). Once these matrices are obtained, we then order their non-$\ddot{0}$ entries by θ_{n+1}. Then F_{n+1} may be constructed from T_{nn} in the following way.

Construction 3. If the kth non-$\ddot{0}$ entry of T_{nn}, relative to θ_{n+1}, is t, then F_{n+1} is obtained by letting its kth row have the value t in its kth place and $\ddot{0}$ elsewhere.

The matrix G_{n+1} may be constructed from H_n by a similar procedure, as follows.

Construction 4. If the kth non-$\ddot{0}$ entry of H_n, relative to θ_{n+1}, is t, then G_{n+1} is obtained by letting its kth row have the value t in its kth place and $\ddot{0}$ elsewhere.

We are now able to state our generalization of Theorem 1.

THEOREM 7. *For any $n \geq 2$, the value matrix of simple n-sequences is given by*

$$M(S_n) = \ddot{A}_{0n} \odot F_n \odot \ddot{A}_{n0}.$$

PROOF. By definition, the i, j entry of $\ddot{A}_{0n} \odot F_n$ is

$$\sum_k \ddot{a}_{ik}^{(0n)} \odot f_{kj}^{(n)}.$$

Since $f_{kj}^{(n)} = \ddot{0}$ for all $k \neq j$, the above expression becomes simply

$$\ddot{a}_{ij}^{(0n)} \odot f_{jj}^{(n)}.$$

Thus, $\ddot{A}_{0n} \odot F_n \odot \ddot{A}_{n0}$ has as its i, j entry

$$\ddot{\sum_k} \ddot{a}_{ik}^{(0n)} \odot f_{kk}^{(n)} \odot \ddot{a}_{kj}^{(n0)},$$

which may be rewritten

$$\ddot{\sum_k} f_{kk}^{(n)} \odot \ddot{a}_{ik}^{(0n)} \odot \ddot{a}_{kj}^{(n0)}.$$

Clearly,

$$\ddot{a}_{ik}^{(0n)} \odot \ddot{a}_{kj}^{(n0)} = \ddot{1}$$

if and only if there is a simple n-sequence from v_i to v_j in N (Theorem 1). Hence, the i, j entry of

$$\ddot{A}_{0n} \odot F_n \odot \ddot{A}_{n0}$$

is the generalized sum of the values of the simple n-sequences from v_i to v_j in N. □

In conclusion, we present an algorithm for finding the value matrix of simple n-sequences of a network. This algorithm is stated in terms of our generalized arithmetic. In specific applications one must, of course, choose appropriate interpretations for the operations of $\odot$ and $\oplus$ and for the identity elements $\ddot{0}$ and $\ddot{1}$. This algorithm makes use of certain steps of Algorithm 1 and employs both binary matrices and value matrices.

Algorithm 2. Let N be a network and θ_n be an arbitrary ordering of the simple n-sequences of N. Then the value matrix $M(S_n)$ of simple n-sequences of N, for $n > 0$, can be obtained by the following procedure.

1. Order the points of N by θ_0 and construct A_{00} and M.
2. Use step 2 of Algorithm 1.

2a. Order the non-$\ddot{0}$ entries of M by θ_1 and construct F_1 using Construction 3. Note that $F_1 = G_1$.

3. Use step 3 of Algorithm 1.

Steps 3a–7 give a recursive procedure for finding A_{nn}, $n = 2, 3, \cdots, p$.

3a. Convert $A_{n-1,n-1}$ to $\ddot{A}_{n-1,n-1}$.

3b. Construct H_{n-1} by the formula

$$H_{n-1} = \ddot{A}_{n-1,n-1} \odot G_{n-1} .$$

3c. Construct $T_{n-1,n-1}$ by the formula

$$T_{n-1,n-1} = F_{n-1} \odot H_{n-1} .$$

4. Use step 4 of Algorithm 1.

4a. Order the non-$\ddot{0}$ entries of $T_{n-1,n-1}$ by θ_n and construct F_n using Construction 3.

4b. Order the non-$\ddot{0}$ entries of H_{n-1} by θ_n and construct G_n using Construction 4.

5. Use step 5 of Algorithm 1.

5a. Convert A_{0n} and A_{n0} to $\ddot{A}_{0n}$ and $\ddot{A}_{n0}$.

6. Use step 6 of Algorithm 1.
7. Use step 7 of Algorithm 1.
8. If $A_{nn} = [0]$ or $n = p$, terminate the procedure. Otherwise, return to step 3a.
9. Construct $M(S_n)$ from the results of steps 5a and 4a by the formula

$$M(S_n) = \ddot{A}_{0n} \odot F_n \odot \ddot{A}_{n0} .$$

Discussion

The algorithms presented here are useful in obtaining information about a variety of structural properties that are based on nonredundant sequences. The matrix S_n of simple n-sequences of D, which is obtained by Algorithm 1, gives the number of n-paths from v_i to v_j and the number of n-cycles containing v_i. Of special interest are the matrices S_{p-1} and S_p, where p is the number of points in D, for one can ascertain from these matrices respectively the number of complete paths and the number of complete cycles in D. From the matrices S_n, for $n \leq p$, one can construct the matrix S of simple sequences which gives the number of paths and cycles of any length in D, the distance matrix which gives the length of a shortest path from v_i to v_j, and the detour matrix which gives the length of a longest path from v_i to v_j. One can also ascertain the length of a shortest cycle and of a longest cycle in D.

Algorithm 2 is applicable when values are associated with the lines of a digraph. If these values are probabilities, the matrix $M(S_n)$, when obtained by use of the appropriate arithmetic, gives the probability of going from v_i to v_j along a path of length n. And if n is the distance from v_i to v_j, the i, j entry is the probability of going from v_i to v_j in the smallest possible number of steps. When the values of lines are costs, one can obtain a matrix $M(S_n)$ whose i, j entry gives the cost of a cheapest simple sequence of length n from v_i to v_j. From the matrices $M(S_n)$, for $n \leq p$, it is possible to ascertain the least cost of any path from v_i to v_j and of any cycle containing v_i. With suitable modifications, similar information can be obtained concerning paths and cycles of maximal cost. Finally, we note that Algorithm 2 can be used when positive or negative signs are assigned to lines. In this case, the i, j entry of $M(S_n)$ indicates the number of positive and of negative paths of length n from v_i to v_j. Such information is useful in the study of the path balance of a structure as developed by Harary, Norman, and Cartwright [2].

The major practical limitation of these algorithms lies in the fact that the rows and columns of the generated matrices correspond to simple sequences of D. Hence, if D has a great many simple sequences of given length, the resulting matrix is large. We have prepared computer programs which can quickly process digraphs containing up to 5,000 simple sequences of a given length. However, the task remains to develop procedures for larger numbers of simple sequences.

REFERENCES

[1] Harary, F. and Norman, R. Z. Some properties of line digraphs. *Rendiconti del Circolo Mathematico di Palermo*, 1960, **9**, 161–168.

[2] Harary, F., Norman, R. Z., and Cartwright, D. *Structural models: An introduction to the theory of directed graphs.* New York: Wiley, 1965.

[3] Hasse, Maria. Über die Behandlung graphentheorischer Probleme unter Verwendung der Matrizenrechnung. *Wiss. Z. Techn. Univer. Dresden*, 1961, **10**, 1313–1316.

[4] Luce, R. D. and Perry, A. D. A method of matrix analysis of group structure. *Psychometrika*, 1949, **14**, 95–116.

[5] Parthasarathy, K. R. Enumeration of paths in digraphs. *Psychometrika*, 1964, **29**, 153–165.

[6] Ross, I. C. and Harary, F. On the determination of redundancies in sociometric chains. *Psychometrika*, 1952, **17**, 195–208.

Manuscript received 5/28/65
Revised manuscript received 9/27/65