Comput. Math. Applic. Vol. 15, No. 4, pp. 337–345, 1988
Printed in Great Britain.

0097-4943/88 $3.00 + 0.00

GRAPH THEORETIC FOUNDATIONS OF PATHFINDER NETWORKS

R. W. SCHVANEVELDT, D. W. DEARHOLT and F. T. DURSO

Computing Research Laboratory, New Mexico State University, Las Cruces, NM 88003, U.S.A.

Abstract—This paper is primarily expository, relating elements of graph theory to a computational theory of psychological similarity (or dissimilarity). A class of networks called Pathfinder networks (PFNETs) is defined. PFNETs are derived from estimates of dissimilarity for pairs of entities. Thus, PFNETs can be used to reveal aspects of the structure inherent in a set of pairwise estimates of dissimilarity. In order to accommodate different assumptions about the nature of the measurement scale (i.e. ordinal, interval, ratio) underlying the data, the Minkowski r-metric (also known as the L norm) is adapted to computing distances in networks. PFNETs are derived from data by: (1) regarding the matrix of dissimilarities as a network adjacency matrix (the DATANET); (2) computing the distance matrix (or r-distance matrix using the Minkowski r-metric) of the DATANET and (3) reducing the DATANET by eliminating each arc that has weight greater than the r-distance between the nodes connected by the arc. PFNET properties of inclusion, relation to minimal spanning trees, and invariance under transformations of data are discussed.

INTRODUCTION

There are several available methods for analyzing similarity or dissimilarity data. Some, such as multidimensional scaling [1–6], assume a continuous, low-dimensional space as the underlying model. Other methods derive from discrete models that yield hierarchical clusters [7], overlapping clusters [8]; tree structures [9–11]; or networks [12–15]. While spatial models have mathematical foundations in geometry, discrete models often derive from graph theory.

The foundations of multidimensional scaling (MDS) have been explored in some depth [16], leading to formal specifications of the assumptions underlying MDS as a model of the psychological representation of stimuli. In recent years, considerable work has gone into the development of discrete models, and the connections between discrete models and graph theory are becoming more apparent (cf. Shepard and Arabie [8]). As representations of mental structure, discrete models offer alternatives that are often closer to psychological theory (e.g. feature, network and categorical theories). In this paper, we discuss network representations and their relationship to dissimilarity data. Pathfinder, an algorithm for deriving networks from dissimilarities, is tied to some fundamental concepts in graph theory. Since much of the discussion revolves around formal properties of networks, a brief review of some basic concepts in graph theory will provide a point of departure.

Graph Theory

Graph theory is the mathematical study of structures consisting of *nodes* with *edges* or *arcs* connecting some pairs of nodes [17–19]. The terminology used in graph theory varies somewhat from one source to another. The presentation here represents a distillation of various sources with adaptations for our purposes.

A *digraph* G is a finite set of nodes (V) and a subset of $V \times V$—the arcs. For example, given a set of nodes $\{1, 2, \ldots, n\}$, the ordered pairs (1, 2), (4, 3), (7, 1) designate arcs from the first to the second node in each pair. A digraph can be displayed by a diagram in which nodes are shown as points, and arcs are indicated by arrows connecting appropriate pairs of points.

The order of the node pair of an arc specifies a *direction* for the arc so that its *initial and terminal nodes* can be distinguished. For example, the arc (3, 2) has node 3 as its initial node and node 2 as its terminal node. In a *symmetric digraph*, for every arc there is another arc that connects the same pair of nodes in the opposite direction. These two arcs are a *symmetric pair*. Symmetric digraphs may be referred to as undirected since a symmetric pair of directed arcs can be represented by an *edge*. A *graph* is a symmetric digraph. In the following discussion, the terms digraph and arc refer to the general case which includes both graphs and digraphs. Some definitions only apply to graphs, and, in such cases, the terms graph and edge refer to symmetric digraphs.

A *walk* is an alternating sequence of nodes and arcs such that the initial node of each arc in the sequence (except the first) is the same as the terminal node of the preceding arc. For example, the arc sequence, (3, 2), (2, 1), (1, 4), specifies a walk, while the arc sequence, (3, 2), (1, 4), (2, 1), does not. A walk can also be specified by the sequence of nodes which it visits. For the example walk specified above, the node sequence is 3, 2, 1, 4. The *length of a walk* corresponds to the number of arcs in the walk. A walk is a *trail* if all arcs in the walk are distinct. A walk is a *path* if all the nodes in the walk are distinct. All paths are trails. An arc is a path of length 1. A *cycle* is a walk with all nodes distinct except the first and last nodes, which are identical. A *connected graph* contains a path (consisting of nodes and edges) between any two nodes.

A *forest* is a graph with no cycles. A *tree* is a connected forest. A tree with n nodes has exactly $n - 1$ edges. In a tree, there is exactly one path between any two nodes.

Arcs may have *weights* (distances or costs) associated with them in which case the digraph is known as a *network*. The digraph corresponding to a network is obtained by deleting the weights. The digraph represents the structure of a network, and the weights associated with arcs in a network provide quantitative information to accompany that structure. The *weight* of arc (i, j) is designated by w_{ij}. In a network, the *weight of a path* can be computed by summing the weights associated with the arcs in the path. The *distance* between two nodes is the minimum weight of paths connecting the nodes. The *minimal spanning tree* [20] of an undirected network consists of a subset of the edges in the network such that the sub-graph is a tree and the sum of the arc weights is minimal over the set of all possible trees.

A *complete graph* has an edge from every node to every other node. A *complete network* is a complete graph with weights associated with the edges.

Various characteristics of digraphs are conveniently represented by matrices. A digraph G can be represented by the *adjacency matrix* A, an $n \times n$ matrix with $a_{ij} = 1$ if G contains the arc (i, j) and $a_{ij} = 0$ otherwise. A network is similarly represented by the *network adjacency matrix* A with $a_{ii} = 0$, $a_{ij} = w_{ij}$, $i \neq j$ if the network contains the arc (i, j), otherwise $a_{ij} = \infty$. The *reachability matrix* of G is the $n \times n$ matrix in which the ijth entry is 1 if there is a path in G from node i to node j and is 0 otherwise. The *distance matrix* D is the $n \times n$ matrix in which d_{ij} is the (minimum) distance from node i to node j in a network. If there is no path from node i to node j, $d_{ij} = \infty$. This distance matrix is not necessarily symmetric, but it will be symmetric if the network consists of edges rather than arcs. An arc in a network is *redundant* if the network obtained by removing the arc yields the same distance matrix as the original network.

Networks as Models

As psychological models, networks entail the assumption that concepts and their relations can be represented by a structure consisting of nodes (concepts) and arcs (relations). Strengths of relations are reflected by arc weights, and the relational meaning of a concept is represented by its connections to other concepts. The use of network models without semantic interpretation of the arcs entails the assumption that the structure in the network corresponds to psychologically meaningful structures. We conjecture that explicit network representations offer the potential of identifying structural aspects of conceptual representation that relate to memory organization, category structure, and other human information processing phenomena.

Less restrictive assumptions are required for using networks as a tool for analyzing data. Networks offer one way, among many, for extracting and representing structure in dissimilarities. The primary assumption is that network representations will simplify the data and will reveal patterns that lead to fruitful interpretations of the dissimilarities.

Network models are frequently encountered in cognitive psychology e.g. Refs [21–25] and artificial intelligence [26–28]. For the most part, however, the networks in these models have been based on intuitions of the researchers. With a few exceptions, which we mention later, networks have not been derived from empirical data.

In contrast, network models have been used on sociometric data for some time [29, 30]. These models characterize relationships among social actors in such social relationships as authority, liking and kinship. Hage and Harary [31] give graph theoretical analyses of several social relations of interest to anthropology. While these applications of graph theory have not been patricularly concerned with dissimilarity data, they have used various kinds of data to determine network structures. The structural analyses available from sociometric network models may prove to be of use in the study of the structure of human knowledge in particular domains. The Pathfinder method of defining networks corresponding to dissimilarity data may also be of use to applications of networks to the analysis of sociometric data.

Hutchinson [12] proposed NETSCAL, an algorithm for constructing networks from dissimilarity data. NETSCAL attempts to identify the arcs that are ordinally necessary given the set of dissimilarities. Also in 1981, we [14] reported some exploratory work on a procedure, Pathfinder, for determining network connections.

Feger and his colleagues [13, 32] have proposed another method known as ordinal network scaling (ONS) which represents rank orders of dissimilarities by a network. Friendly [33, 34] and Fillenbaum and Rapoport [35] investigated some direct methods for establishing network structures on the basis of empirical data. Friendly used a threshold on ratings of similarity to determine which edges to include in a network. Fillenbaum and Rapoport asked people to create networks directly by drawing them. All of these techniques hold the promise of placing network representations on a firmer empirical foundation. It would be of value, however, to establish formal relationships between empirically derived networks and graph theory.

DISSIMILARITIES AND GRAPH THEORY

How are dissimilarities related to graph theory? A common form for representing pairwise dissimilarities is an $n \times n$ matrix P, where p_{ij} is the dissimilarity of entity i and entity j. Then p_{ii} represents the dissimilarity of an entity with itself which is usually assumed to be zero. If the dissimilarity estimates are symmetrical ($p_{ij} = p_{ji}$, for all i, j), the matrix will be symmetric about the major diagonal. Dissimilarities need not be symmetrical, however, and no such constraint need be imposed on matrices associated with networks. Of the matrices we have considered, we might regard the dissimilarity matrix either as the distance matrix of a network or as the adjacency matrix of a network. Either of these could provide a fruitful connection with graph theory so that algorithms on networks can be used to extract graph and digraph structures from dissimilarities. We will explore both of these alternatives.

Dissimilarities as Distances

If we regard the dissimilarity matrix as the distance matrix of a network, then our problem is to determine what network could have produced those distances. This is essentially the basis of Hutchinson's NETSCAL algorithm. We will follow his analysis which is based on two theorems relating distance matrices to networks. A matrix D is *realizable* as a network if D is the distance matrix of some network.

Theorem 1 [36]

An $n \times n$ matrix D is realizable as a network if and only if

(a) identity: $d_{ii} = 0$;
(b) positivity: $d_{ij} > 0, \quad i \neq j$ and
(c) triangle inequality: $d_{ij} \leq d_{ik} + d_{kj}, \quad$ for all i, j, k.

A matrix D satisfying identity, positivity, and the triangle inequality does not necessarily correspond to a unique network. The complete network with $w_{ij} = d_{ij}$, $i \neq j$ is always one realization of D. A given realization of D, however, may contain redundant arcs which can be deleted from the network without changing any distances. An *irreducible network* is one with no redundant arcs. Theorem 2 specifies the necessary and sufficient conditions for an irreducible realization of D.

Theorem 2 [37]

Given a distance matrix D, an arc (i, j) is an arc in the (unique) irreducible realization of D if and only if

$$d_{ij} < \min(d_{ik} + d_{kj}), \quad i \neq j, \quad k \neq i, \quad k \neq j.$$

Another way of stating this result is that the arc (i, j) is not in the irreducible network if and only if that network contains an alternative path connecting nodes i and j with weight equal to d_{ij}. (The weight of the alternative path cannot be less than d_{ij} because of the triangle inequality.)

When we begin with a set of numbers that has been obtained from human judgement, we should be particularly concerned about the measurement scale (e.g. ordinal, interval, ratio) underlying the numbers [38]. The level of measurement that holds for data determines which properties of numbers we can ascribe to the data. An ordinal level of measurement means that the data values are properly ordered, but comparing differences may not be meaningful. A ratio scale (which is typical of physical measurement) allows meaningful statements about the ratios of values so that we can say that x has twice as much of some quantity as y. It is useful to consider allowable transformations with the various levels of measurement. With ordinal measurement, any nondecreasing function is allowed since the values will retain their order. With ratio measurement, the only allowable transformation is multiplication by a positive constant (i.e. a change of unit).

Realizing that dissimilarity data are usually not measured on a ratio scale which is assumed for Theorem 2, Hutchinson [12] based the NETSCAL algorithm on a corollary of Theorem 2 which provides ordinally sufficient but not necessary conditions for the presence of arcs.

Corollary 1 [12]

Given a distance matrix D, the arc (i, j) is present in the (unique) irreducible realization of D if

$$d_{ij} \leqslant \min[\max(d_{ik}, d_{kj})], \quad i \neq j, \quad k \neq i, \quad k \neq j.$$

The major problem with this approach is the assumption that empirically obtained dissimilarity matrices meet the conditions for a distance matrix to be realizable as a network. Positivity can usually be satisfied by assumption and constraints on dissimilarities. However, the triangle inequality will not necessarily hold for dissimilarities. Of course, Theorems 1 and 2 are concerned with the relations between networks and their distance matrices when the relation is exact. Our problem in working with dissimilarities is somewhat different. Aside from the fact that dissimilarities usually contain error (to which we will return later), psychological judgments may show systematic violations of the triangle inequality [39]. For one thing, entities may be related to one another in different ways. To use Tversky's example, Jamaica is similar to Cuba and Cuba is similar to Russia, but Jamaica is not at all similar to Russia. Also, psychological judgment may not be as transitive as logic would suggest. For example, while people may judge that successive items in the list (forks, silverware, furnishings, manufactured goods and things) are related to one another, the degree of relatedness may decrease rapidly for pairs further separated in the list. (How closely related are forks and things?) Perhaps regarding dissimilarities as arc weights will provide a way around the dubious assumption that dissimilarities conform to the triangle inequality.

Dissimilarities as Arc Weights

If we regard the dissimilarity matrix P as an adjacency matrix A such that $a_{ij} = p_{ij}$, the corresponding network is complete if all dissimilarities are finite, certainly an accurate representation of the dissimilarities, but not very informative. However, because the network represented by A is a realization of its distance matrix D, the positivity and triangle inequality conditions of Theorem 1 are satisfied by D, and Theorem 2 defines the irreducible realization of D. The irreducible realization of D may consist of fewer arcs than A, and thus may be more informative than A about the latent structure of the dissimilarities. However, as mentioned earlier, Theorem 2, requires measurement on a ratio scale which is usually not true of dissimilarities. One solution to this problem is a generalization of the definition of distance in a network that will accommodate ordinal as well as ratio scales of measurement.

Distances in Networks

Usually, in graph theory, the distance between nodes i and j is the minimum weight of all possible paths from i to j, $i \neq j$ where the weight of a path is the sum of the wieghts of the arcs in the path. From the perspective of measurement scales underlying dissimilarity data, it would be useful to define a distance function that will permit computations of distances in networks with different assumptions about the level of measurement associated with the dissimilarities. Such a distance function should preserve ordinal relationships between arc weights and path weights for all permissible transformations of the dissimilarities with different assumptions about the level of measurement associated with the dissimilarities. Then Theorem 2 would identify the same arcs in the irreducible realization of the distance matrix for all premissible transformations on the dissimilarities.

A distance function with the required qualities can be defined by adapting the Minkowski distance measure to computing distances over paths in networks. This distance function replaces addition with the r-metric computation so that $x + y$ is replaced by $(x^r + y^r)^{1/r}$, $r \geqslant 1$. Given a path P consisting of k arcs, the weight of P, $w(P)$ becomes

$$w(P) = \left[\sum_{i=1}^{k} w_i^r \right]^{1/r}, \quad 1 \leqslant r \leqslant \infty.$$

Note that with $r = 1$, the r-metric function corresponds to simple addition. With $r = \infty$, the r-metric is the maximum function. In fact,

$$\lim_{r \to \infty} (x^r + y^r)^{1/r} = \max(x, y).$$

Thus, with $r = \infty$, computing path distances with the Minkowski r-metric only requires maximum and minimum operations which are order preserving and, therefore, appropriate for ordinal scale measurement. In particular, the ordinal relationships of path weights will be preserved for any nondecreasing transformation of the arc weights (dissimilarities).

It can be easily shown that the Minkowski r-metric satisfies the requirements of a path algebra as defined by Carre [17].

An attractive property of the Minkowski r-metric is that a single weight can be associated with a path regardless of segmentation. Given an exhaustive set of path segments, S associated with path P (i.e. S is a decomposition of path P into sub-paths.)

$$w(P) = \left[\sum_{s \in S} [W(s)]^r \right]^{1/r}.$$

The use of the r-metric to compute path weights requires the assumption that the arcs in a path represent independent contributions to the total weight of the path. Increasing the value of r increases the relative contribution of the larger weights in a path. Following a suggestion by Cross (1965), cited in Coombs *et al.* [40], r may be interpreted as a parameter of component weight. With $r = 1$, all components (arcs in a path) have equal weight in determining the weight of a path. As r increases, the components with greater magnitude receive greater weight until, in the limit, only the largest component (arc) determines the weight of a path. One psychological interpretation of larger values of r is that the perceived dissimilarity between entities is determined by the dissimilarity of the most dissimilar relations connecting the entities.

We can generalize Theorem 1 in terms of the r-metric definition of distance.

*Theorem 1^r^*

Let $1 \leqslant r \leqslant \infty$. An $n \times n$ matrix $D = [d_{ij}]$ is r-realizable as a network if and only if

(a) identity: $d_{ii} = 0$;
(b) positivity: $d_{ij} > 0, \quad i \neq j$ and
(c) the r-metric inequality: $d_{ik} \leqslant (d_{ij}^r + d_{jk}^r)^{1/r}, \quad r \geqslant 1,$ for all i, j, k.

The proof follows that of Hakimi and Yau [36] exactly.

While Theorem 1 generalizes readily to the r-metric definition of distance, Theorem 2 does not. Distance matrices computed with $r = \infty$ present particular difficulties. If we remove all arcs that do not satisfy the inequality associated with $r = \infty$: $d_{ij} < \min[\max(d_{ik}, d_{kj})]$, $i \neq j$, $i \neq k$, $j \neq k$, we cannot, in general, reproduce the distance matrix with the resulting network.

Thus the r-metric generalization of network distance complicates the notion of redundant arcs. There are actually two types of redundant arcs: (a) arcs with weight greater than the weight of an alternative path and (b) arcs with weight equal to the weight of the minimum weight alternative path. Type (a) redundant arcs can be eliminated without difficulty and the resulting distance matrix is unchanged regardless of the value of the r-metric. Type (b) redundant arcs, on the other hand, cannot be eliminated, in general, without changing the distance matrix associated with the network.

With $r = \infty$, type (b) redundant arcs occur in sets of two or more arcs with equal weight. Eliminating all of the arcs in a redundant set changes the distance matrix. However, subsets of the arcs in a redundant set can be eliminated without changing the distance matrix. The problem is that various subsets can be eliminated and there is no canonical way of selecting one subset over another. One direct solution to this problem is to define a reduction of a network that excludes type (a) redundant arcs while including type (b) redundant arcs.

A *reduction of a network* G is a network G' such that the arcs in G' are a subset of the arcs in G and G and G' have the same r-distance matrix (D). The *triangular reduction* of a network G with r-distance matrix $D = [d_{ij}]$ is obtained by removing from G every arc (i, j) with weight greater than d_{ij}. The following observation gives necessary and sufficient conditions for arcs to be in the triangular reduction of G.

> Given a network G with adjacency matrix $A = [a_{ij}]$ and r-distance matrix $D = [d_{ij}]$, an arc (i, j) in G is an arc in the triangular reduction of G if and only if $a_{ij} \neq \infty$ and $d_{ij} = a_{ij}$, $i \neq j$.

The triangular reduction is a reduction. The triangular reduction may contain arcs whose weights are equal to the weight of a minimum weight alternative path while the irreducible realization in the case of $r = 1$ will not. Both the triangular reduction and the irreducible realization exclude arcs with weights greater than the weight of a minimum weight alternative path. The triangular reduction will also be the irreducible realization if there are no arcs in G with weight equal to the minimum weight of alternative paths. One characteristic of the triangular reduction of G is that it minimizes the length (number of arcs) of the minimum weight paths.

The definition of the triangular reduction provides a basis for removing arcs from the network with adjacency matrix $A = P$, the dissimilarity matrix. The Pathfinder algorithm is a realization of this network reduction criterion.

THE PATHFINDER ALGORITHM

Let us call a network resulting from the application of the Pathfinder algorithm a PFNET of the original network. The essential idea behind Pathfinder is that dissimilarities between entities should be represented as arcs in a PFNET if the resulting arcs form the minimum weight paths given the set of dissimilarity estimates. In fact, the definition of Pathfinder can be stated quite simply: Given a network G defined by dissimilarities, the PFNET(r) is the triangular reduction of G with r-distance matrix D.

The derivation of Pathfinder networks requires computing the distance matrix of a complete (or nearly complete) network. This computation has time complexity of $O(n^3)$. While this complexity is prohibitive for rapid computation on large networks, it is quite manageable for occasional derivations of networks with hundreds of nodes. Several potential applications of Pathfinder require analysis of problems of this size.

Because different values of r result in different weights of paths, Pathfinder can produce several different PFNETs. We now turn to an examination of some of the Pathfinder PFNETs and their relations to one another.

The r-metric yields systematic variation in path weights as r varies over the allowable range. This variation is expressed in the following theorem in terms of the r-distance matrix D computed on the network G defined by the dissimilarity data.

Theorem 3

Given a network G and r-distances d_{r_1} and d_{r_2} computed on G:

$$d_{r_1} \leqslant d_{r_2} \quad \text{if and only if} \quad r_1 \geqslant r_2.$$

The proof of Theorem 3 makes use of the following lemma.

Lemma 1.

Given $w_1, \ldots, w_k$, $w_i \geqslant 0$ for all i:

$$\left[\sum_{i=1}^{k} w_i^{r_1}\right]^{1/r_1} \leqslant \left[\sum_{i=1}^{k} w_i^{r_2}\right]^{1/r_2} \quad \text{if and only if} \quad r_1 \geqslant r_2.$$

This lemma is well known in the literature of harmonic analysis and functional analysis. A concise proof is presented by Edwards [41, p. 29].

Theorem 3 is immediate from Lemma 1. The lemma shows that the theorem holds for paths, and upon taking minimums, it holds for distance functions.

Inclusion Relationships among PFNETs

A network G' is *included in* in a network G if G and G' have the same nodes and the arcs in G' are a subset of the arcs in G. We also say that network G *includes* network G'. The following immediate consequence of Theorem 3 establishes an inclusion relationship among various PFNETs.

Corollary 2

PFNET(r_1) is included in PFNET(r_2) if and only if $r_1 \geqslant r_2$.

The definition gives the criterion for including an arc in PFNET(r), the triangular reduction of the network defined by the dissimilarities with r-distance matrix D. Including arc (i, j) in PFNET(r) requires that the dissimilarity (a_{ij}) be equal to the ijth entry of $D = [d_{ij}]$. From Theorem 3 we know that $d_{r1} \leqslant d_{r2}$, $r_1 \geqslant r_2$. Since decreasing r can only increase d_r and a_{ij} is an upper bound on d_{ij}, every arc (i, j) in PFNET (r_1) is also an arc in PFNET(r_2), $r_1 \geqslant r_2$.

A family of PFNETs can be generated by variations in r. As a result of the inclusion relationship, PFNET(r)'s exhibit a monotonic decrease in the number of arcs as r increases. Thus, we can select a particular PFNET in the family of PFNETs by specifying a value of r between one and infinity.

The *minimally connected PFNET* is PFNET(∞). The PFNET(∞) has the fewest arcs of any PFNET for a particular set of data. With symmetrical dissimilarity data, the PFNET(∞) is the union of all minimal spanning trees [20, 42] for the network defined by the dissimilarities. The PFNET(∞) will be the unique minimal spanning tree when there is such a unique tree. Certain patterns of ties in the dissimilarity data may result in there being more than one tree in which case the PFNET(∞) will include all edges that are in any minimal spanning tree. The PFNET(∞) represents the simplest unique PFNET for a given set of dissimilarities. Figure 1(b) shows the PFNET(∞) for the dissimilarity data in Fig. 1(a). The PFNET(∞), in this case, is a tree (no cycles), and it is the minimal spanning tree for the complete network shown in Fig. 1(a).

Using different r values to compute path weight will usually produce different PFNETs. For example, PFNET(1) is the result of using the usual sum of the arc weights in a path to define the path weight function. PFNET(1) includes all of the arcs in the PFNET(∞), but PFNET(1) will usually have additional arcs as well. Figure 1(c) shows PFNET(1) for the dissimilarity data in Fig. 1(a). In this case, PFNET(1) has two additional arcs over the PFNET(∞), and the additional arcs necessarily introduce cycles.

Levels of Measurement

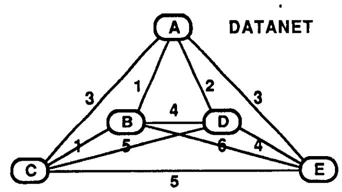
Although variation in the r parameter has the value of allowing control over the number of arcs in the PFNET, assumptions about the dissimilarity estimates should influence the choice of values for r. In particular, the measurement scale underlying the dissimilarity estimates places constraints on values of r because different PFNET structures can result from applying Pathfinder to transformed data. In would be desirable to select values of r so that the same arcs would be present in the PFNETs generated from all permissible transformations of the dissimilarity estimates.

With measurement on a ratio scale [38], the only allowable transformations involve multiplication by a positive constant (i.e. a change of unit). Pathfinder will preserve the PFNET structure (i.e. have exactly the same arcs) under multiplication of the dissimilarity estimates by a positive constant for all values of r. Thus, with ratio-level measurement, any value of r can be used, and the selection of r can be determined by the desired number of arcs in the PFNET or other criteria.

(a) Data (dissimilarities)

A = P =

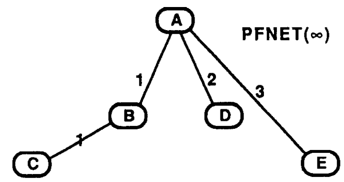
	A	B	C	D	E
A	0	1	3	2	3
B	1	0	1	4	6
C	3	1	0	5	5
D	2	4	5	0	4
E	3	6	5	4	0



(b) $r = \infty$

D =

	A	B	C	D	E
A	0	1	1	2	3
B	1	0	1	2	3
C	1	1	0	2	3
D	2	2	2	0	3
E	3	3	3	3	0



(c) $r = 1$

D =

	A	B	C	D	E
A	0	1	2	2	3
B	1	0	1	3	4
C	2	1	0	4	5
D	2	3	4	0	4
E	3	4	5	4	0

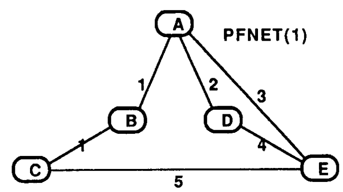


Fig. 1. (a) The adjacency (or dissimilarity) matrix corresponding to the data network. (b) The distance matrix and the PFNET for $r = \infty$. The minimal spanning tree. (c) The distance matrix and the PFNET for $r = 1$.

With psychological measurement, we are often only willing to assume that scale values represent ordinal information, and, as a result, the "true" scale values may be any nondecreasing function of the actual values in the data. With such ordinal level measurement [38], Pathfinder will provide a unique PFNET structure only for $r = \infty$. That is, the same arcs will be present in PFNET(∞) derived from any nondecreasing transformation of a particular set of dissimilarities. Thus, the PFNET(∞) is a unique structure for levels of measurement ranging from ordinal through interval to ratio. It is the only unique structure with ordinal measurement.

Applications

We and others have been investigating applications of Pathfinder to problems in cognitive modeling, knowledge representation, knowledge elicitation, and user-computer interface design. Details on these efforts can be found in the following Refs [15, 43–49].

Acknowledgements—This work was accomplished under the sponsorship of the Computing Research Laboratory and the National Science Foundation (Grant No. IST-8506706). F. Durso is at the Department of Psychology, University of Oklahoma. Appreciation is expressed to Tim Goldsmith, Frank Harary, Jim McDonald, Keith Phillips and Mark Wells for their comments and suggestions.

REFERENCES

1. R. N. Shepard, Analysis of proximities: multidimensional scaling with an unknown distance function. I. *Psychometrika* **27,** 125–140 (1962).
2. R. N. Shepard, Analysis of proximities: multidimensional scaling with an unknown distance function. II. *Psychometrika* **27,** 219–246 (1962).
3. R. N. Shepard, Analysis of proximities as a technique for the study of information processing in man. *Human Factors* **5,** 33–48 (1963).
4. R. N. Shepard, Representation of structure in similarity data: problems and prospects. *Psychometrika* **39,** 373–421 (1974).
5. J. B. Kruskal, Nonmetric multidimensional scaling: a numerical method. *Psychometrika* **29,** 115–129 (1964).
6. J. B. Kruskal, Multidimensional scaling and other methods for discovering structure. *Statistical Methods for Digital Computers* (Eds Enstein, Ralston and Wilf). Wiley, New York (1977).

7. S. C. Johnson, Hierarchical clustering schemes. *Psychometrika* **32,** 241–254 (1967).
8. R. N. Shepard and P. Arabie, Additive clustering: representation of similarities as combinations of discrete overlapping properties. *Psychol. Rev.* **86,** 87–123 (1979).
9. K. A. Butler and J. E. Corter, The use of psychometric tools for knowledge acquisition: a case study. In *Artificial Intelligence and Statistics.* (Ed. W. Gale). Addison-Wesley, Reading (1986).
10. J. P. Cunningham, Free trees and bidirectional trees as representations of psychological distance. *J Math. Psychol.* **17,** 165–188 (1978).
11. S. Sattath and A. Tversky, Additive similarity trees. *Psychometrika* **42,** 319–345 (1977).
12. J. W. Hutchinson, Network representations of psychological relations. Unpublished Ph.D. dissertation, Stanford University (1981).
13. H. Feger and W. Bien, Network unfolding. *Social Networks* **4,** 257–283 (1982).
14. R. W. Schvaneveldt and F. T. Durso, General semantic networks. Paper presented at the annual meetings of the Psychonomic Society, Philadephia (1981).
15. R. W. Schvaneveldt, F. T. Durso, and D. W. Dearholt, Pathfinder: Scaling with Network Structures. Memorandum in Computer and Cognitive Science, MCCS-85-9, Computing Research Laboratory, New Mexico State University (1985).
16. R. Beals, D. H. Krantz and A. Tversky, Foundations of multidimensional scaling. *Psychol. Rev.* **75,** 127–142 (1968).
17. B. A. Carre, *Graphs and Networks.* Clarendon Press, Oxford (1979).
18. N. Christofides, *Graph Theory: An Algorithmic Approach.* Academic Press, New York (1975).
19. F. Harary, *Graph Theory.* Addison-Wesley, Reading, Mass. (1969).
20. J. B. Kruskal, On the shortest spanning subtree of a graph and the traveling salesman problem. *Proc. Am. Math Soc* **7,** 48–50 (1956).
21. J. R. Anderson, *The Architecture of Cognition.* Harvard University Press, Cambridge, Mass. (1983).
22. A. M. Collins and E. F. Loftus, A spreading activation theory of semantic processing. *Psychol. Rev.* **82,** 407–428 (1975).
23. A. M. Collins and M. R. Quillian, Retrieval time from semantic memory. *J. Verbal Learning Verbal Behavior* **8,** 240–247 (1969).
24. D. E. Meyer and R. W. Schvaneveldt, Meaning, memory structure and mental processes. *Science* **192,** 27–33 (1976).
25. R. W. Schvaneveldt and D. E. Meyer, Retrieval and comparison processes in semantic memory. In *Attention and Performance IV* (Ed. S. Kornblum). Academic Press, New York (1973).
26. M. R. Quillian, The teachable language comprehender: A simulation program and theory of language. *Communs. ACM* **12,** 459–476 (1969).
27. M. Minsky, A framework for representing knowledge. In *The Psychology of Computer Vision* (Ed. P. Winton). McGraw Hill, New York (1975).
28. R. C. Schank and R. P. Abelson, *Scripts, plans, goals, and understanding.* Erlbaum, Hillsdale, New Jersey (1977).
29. F. Harary, R. Z. Norman and D. Cartwright, *Structural Models: An Introduction to the Theory of Directed Graphs.* Wiley, New York (1965).
30. D. Knoke and J. H. Kuklinski, *Network Analysis.* Sage Publications, Beverly Hills, Calif. (1982).
31. P. Hage and F. Harary, *Structural Models in Anthropology.* Cambridge University Press (1983).
32. U. Droge and H. Feger, Ordinal network scaling. Paper presented at the joint meeting of the Classification Society and the Psychometric Society, Paris (1983).
33. M. Friendly, In search of the M-gram: the structure of organization in free recall. *Cogn. Psychol.* **9,** 188–249 (1979).
34. M. Friendly, Methods for finding graphic representations of associative memory structures. In *Memory Organization and Structure* (Ed. C. R. Puff). Academic Press, New York (1979).
35. S. Fillenbaum and A. Rapoport, *Structures in the Subjective Lexicon.* Academic Press, New York (1971).
36. S. L. Hakimi and S. S. Yau, Distance matrix of a graph and its realizability. *Q. Jl Appl. Math.* **22,** 305–317 (1964).
37. A. J. Goldman Realizing the distance matrix of a graph. *J. Res. natn Bureau Stand.* **70B,** 153–154 (1966).
38. S. S. Stevens, Mathematics, measurement, and psychophysics. In *Handbook of Experimental Psychology* (Ed. S. S. Stevens). Wiley, New York (1951).
39. A. Tversky, Features of similarity. *Psychol. Rev.* **84,** 327–352 (1977).
40. C. H. Coombs, R. M. Dawes and A. Tversky, *Mathematical Psychology: An Elementary Introduction.* Prentice-Hall, Englewood Cliffs, New Jersey (1970).
41. R. E. Edwards, *Fourier Series: A Modern Introduction*, Vol. 1, 2nd edn. Springer, New York (1979).
42. D. W. Dearholt, R. W. Schvaneveldt and F. T. Durso, Properties of networks derived from proximities. Memorandum in Computer and Cognitive Science, MCCS-85-14, Computing Research Laboratory, New Mexico State University (1985).
43. N. M. Cooke, F. T. Durso and R. W. Schvaneveldt, Recall and measures of memory organization. *J. Exp. Psychol. Learning Memory Cognition* **12,** 518–549 (1986).
44. N. M. Cooke and J. E. McDonald, A formal methodology for acquiring and representing expert knowledge. *Proc. of the IEEE* **110,** 1422–1430 (1986).
45. J. E. McDonald, D. W. Dearholt, K. R. Paap and R. W. Schvaneveldt, A formal interface design methodology based on user knowledge. *Proc. CHI'86*, pp. 285–290 (1986).
46. J. E. McDonald and R. W. Schvaneveldt, The application of user knowledge to interface design. In *Cognitive Science and its Application to Human-Computer Interactions* (Ed. R. Guindon) Erlbaum, Hillsdale, New Jersey (in press).
47. R. J. Roske-Hofstrand and K. R. Paap, Cognitive networks as a guide to menu organization: an application in the automated cockpit. *Ergonomics* **29,** 1301–1311 (1986).
48. R. Schvaneveldt, F. Durso, T. Goldsmith, T. Breen, N. Cooke, R. Tucker and J. DeMaio, Measuring the structure of expertise. *Int. J. Man-Machine Stud.* **23,** 699–728 (1985).
49. K. Snyder, K. Paap, J. Lewis, J. Rotella, A. Happ, L. Malcus and J. Dyck, Using cognitive networks to create menus. IBM Technical Report No. TR 54.405, Boca Raton, Fla (1985).