NETWORK STRUCTURES IN PROXIMITY DATA

Roger W. Schvaneveldt
Francis T. Durso
Donald W. Dearholt

I. Introduction

Proximity data are commonplace in the social and behavioral sciences. Judgments of similarity, relatedness, or association between entities are frequently used in the study of human cognition.[1] Investigations of social processes often make use of proximity measures such as liking between pairs of individuals and frequency of communication between individuals or groups of individuals. Proximities can also be obtained from measures of co-occurrence, sequential dependency, correlation, and distance.

This ubiquity of proximity data has encouraged the development of many methods for characterizing the underlying structure in sets of proximities. Some methods, such as multidimensional scaling (Shepard, 1962a, 1962b; Kruskal, 1964, 1977), assume a continuous, low-dimensional space as the underlying model. Spatial models generally represent

[1] Similarity, relatedness, and psychological distance are closely related concepts indicating the degree to which things belong together psychologically. Proximity is a general term that represents these concepts as well as other measurements, both subjective and objective, of the relationship between pairs of entities. In this chapter, we use the term *proximity* to refer to such measurements. In the techniques we propose, the measurements have the direction of distances (or distance estimates) so that small values represent similarity, relatedness, or nearness, and large values represent dissimilarity, lack of relatedness, or distance.

entities as points in space and relations between entities are captured in distances between entities in that space. The dimensions of the space often reflect important dimensions of variation in the proximity data. Other methods derive from discrete models that yield hierarchical clusters (Johnson, 1967), overlapping clusters (Shepard & Arabie, 1979), tree structures (Butler & Corter, 1986; Cunningham, 1978; Sattath & Tversky, 1977), or networks (Hutchinson, 1981; Feger & Bien, 1982; Schvaneveldt & Durso, 1981; Schvaneveldt, Dearholt, & Durso, 1988). Discrete models generally represent entities as nodes in networks and relations between entities as links connecting nodes. The patterns of connections among nodes in networks often reflect clustering and other structures in the proximity data. Whereas 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 (Beals, Krantz, & Tversky, 1968), leading to formal specifications of the assumptions underlying MDS as a model of the psychological representation of stimuli. Considerable work has gone into the development of discrete models, and the connections between discrete models and graph theory are becoming more apparent (Shepard & Arabie, 1979; Schvaneveldt *et al.,* 1988). As representations of mental structure, discrete models offer alternatives to spatial models that are often more closely identified with psychological theory, particularly network-based models.

In this chapter, we discuss network representations and their relationship to proximity data. Pathfinder, a definition of a class of networks derived from proximity data, is tied to some fundamental concepts in graph theory, and illustrations of the application of Pathfinder networks to a variety of data are presented. Since much of the discussion revolves around graphs and networks, we briefly review some basic concepts.

A. Graph Theory

Graph theory is the mathematical study of structures consisting of *nodes* with *links* connecting some pairs of nodes (Carre, 1979; Christofides, 1975; Harary, 1969). Terminology in graph theory varies somewhat from one source to another. Our terms represent a distillation of various sources with adaptations to our purposes.

A *graph G* consists of *nodes* and *links.* The nodes are a finite set, such as $\{1, 2, \ldots, n\}$, and the links are a subset of the set of all node pairs. For example, the node pairs (1, 2), (4, 3), (7, 1) designate links between the first and the second node in each pair. The nodes connected by a link are known as *endpoints* of the link. A link is *incident* to a node if the node is an endpoint of the link. The *degree of a node* is the number of links incident to the node. A graph can be displayed by a diagram in which nodes are shown as points and links are indicated by lines or arrows connecting appropriate pairs of points.

A graph may be either directed or undirected. A *directed graph* (sometimes referred to as a *digraph*) has directed links (or *arcs*). The order of the nodes in a pair designating an arc specifies a direction for the arc, which is regarded as going from the first (or *initial*) node to the second (or *terminal*) node. In diagrams of directed graphs, arcs are represented as arrows extending from the initial node to the terminal node. An *undirected graph* has undirected links (or *edges*). The nodes in a pair designating an edge are regarded as unordered. In diagrams of undirected graphs, edges are represented as lines connecting appropriate nodes. In our usage, the terms graph and link refer to the general case, which includes both directed and undirected graphs.

A *walk* is an alternating sequence of nodes and links such that each link in the sequence connects the nodes that precede and follow it in the sequence. For example, given nodes {1, 2, 3, 4}, the sequence, 3, (3,2), 2, (2,1), 1, (1,4), 4, specifies a walk, whereas the sequence, 3, (3,2), 2, (1,4), 4, (2,1), 1 does not. A walk can be specified by the sequence of nodes that it visits, in which case the existence of the appropriate links is assumed. For the exemplary walk specified above, the node sequence is 3,2,1,4. The *length of a walk* corresponds to the number of links in the walk. A walk is a *path* if all the nodes in the walk are distinct. A link 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 between any two nodes. A *tree* is a connected graph with no cycles. An undirected tree with n nodes has exactly $n - 1$ edges, and it contains exactly one path between any two nodes. A *complete graph* has all possible links.

Links may have positive real numbers (weights, distances, or costs) associated with them in which case the graph is known as a *network*. The graph corresponding to a network is obtained by deleting the weights. The graph represents the structure of a network, and the weights associated with links in a network provide quantitative information to accompany that structure. The *weight* of link *(ij)* is designated by w_{ij}. A graph may be regarded as a network with all link weights equal to one (1). In a network, the *weight of a path* is the sum of the weights associated with the links in the path. A *geodesic* is a minimum weight path connecting two nodes. The *distance* between two nodes is the weight of a geodesic connecting the nodes. The *minimal spanning tree* (Kruskal, 1956) of an undirected network consists of a subset of the edges in the network such that the subgraph is a tree and the sum of the link weights is minimal over the set of all possible trees.

Various characteristics of graphs are conveniently represented by matrices. A graph G can be represented by the *adjacency matrix A*, the $n \times n$ matrix with $a_{ij} = 1$ if G contains the link *(ij)* 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 link (i, j), otherwise $a_{ij} = \infty$. The *reachability matrix* of G is the $n \times n$ matrix in which the ij^{th} entry is 1 if there is a path in G from node i to node j and is 0 otherwise. The *distance matrix D* of a network 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 (a disconnected network), $d_{ij} = \infty$. The distance matrix of a graph contains the (minimum) number of links between pairs of nodes. The distance matrix is not necessarily symmetric, but it will be symmetric if the network consists of undirected links. A link in a network is *redundant* if the network obtained by removing the link yields the same distance matrix as the original network.

B. 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 links (relations). Strengths of relations are reflected by link weights, and the intensional meaning of a concept is determined by its connections to other concepts. As discussed in the later section on applications, networks can be used to model heterogeneous sets of relations on concepts, in which case we assume that the links have a semantic interpretation such as those found in semantic networks (e.g., Quillian, 1969; Collins & Loftus, 1975; Meyer & Schvaneveldt, 1976). The use of network models without interpretation of the links entails the assumption that the structure in the network corresponds to psychologically meaningful relations. Alternatively, we might assume that the network identifies salient associations between concepts.

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 knowledge-based phenomena. We have begun to explore this conjecture and review some of our work in this area in the applications section.

Less restrictive assumptions are required for using networks as a descriptive tool for analyzing proximity data. Networks offer one way among many for extracting and representing structure in proximities. The primary requirement for description is that network representations reveal patterns in data that lead to fruitful interpretations.

Network models have been used on sociometric data for some time (Harary, Norman, & Cartwright, 1965; Knoke & Kuklinski, 1982). These models characterize relationships among social actors in such social relationships as authority, liking, and kinship. Hage and Harary (1983) give graph-theoretic analyses of several social relations of interest to anthropology. Although these applications of graph theory have not been particularly concerned with proximity 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 proximity data may also be of use in applications of networks to the analysis of sociometric data.

C. Network Representations

In applications of networks, the nodes usually represent entities, and the links represent pairwise relations between the entities. Because a set of nodes can be connected by links in many possible ways, a wide variety of structures can be represented by graphs.

Trees are the basis of such psychometric methods as hierarchical cluster analysis (Johnson, 1967), weighted free trees (Cunningham, 1978), and additive similarity trees (Sattath & Tversky, 1977). All of these methods require estimates of pairwise proximities and yield some form of tree structure corresponding to the data.

Hierarchical cluster analysis provides a set of nested (hierarchical) groupings of the entities intended to correspond to meaningful categories. Different hierarchical clustering methods use different definitions of the proximity between a category (once formed) and the other entities and categories. The *single link* method uses the minimum of the proximities between the entities in a category and the entities in other categories. The *complete link* method uses the maximum proximity. Another variation uses the average proximity between entities in different categories. The value of hierarchical cluster analysis lies in its potential for revealing the underlying categorical structure for a set of entities. One problem often encountered in uses of cluster analysis stems from the necessity for clusters to be nested, which means that an entity can only belong to certain clusters.

Additive clustering (Shepard & Arabie, 1979) is a method for producing overlapping clusters so that an entity may belong to more than one cluster. The clusters are not necessarily nested, so that nonhierarchical structures can be revealed. Such a representation violates the constraints on a tree structure and thus corresponds to a general graph. The theory underlying additive clustering assumes that the entities have associated sets of features, and the clusters correspond to shared features among the entities. The value in the method lies in its ability to suggest these underlying features.

Networks have also played an important role in theoretical work on memory structure and knowledge representation (e.g., Anderson, 1983; Collins & Loftus, 1975; Meyer & Schvaneveldt, 1976; Quillian, 1969). In practice, the actual networks employed have been based largely on logical analysis or the intuitions of theorists. There are some notable exceptions, however. Fillenbaum and Rapoport (1971) asked people to construct networks by indicating which pairs of items should be connected. This method assumes that people have introspective access to the information required to characterize the network structure. This is a rather strong assumption, and more indirect methods for identifying networks would be desirable.

Friendly (1977, 1979) produced networks representing associative memory structures by using a threshold on the proximities between items (nodes) in free recall to determine which nodes to connect. Those pairs of items that were "closer" than the threshold were connected in the resulting network. Friendly's method does not require people to have explicit knowledge of network structures, but the use of a threshold can be problematic in that it does not take the relative relations between nodes into account. In contrast, Pathfinder networks, as we shall show, determine link membership by the relations between the possible paths connecting nodes.

Hutchinson (1981) proposed NETSCAL, an algorithm for constructing networks from proximity data. NETSCAL attempts to identify the links that are ordinally necessary given the set of proximities. Also in 1981, we (Schvaneveldt & Durso, 1981) reported some exploratory work on Pathfinder networks. As it turns out, Pathfinder defines a family of networks for a given set of proximity data. One of the networks in this family is identical to the one generated by NETSCAL.

Feger and his colleagues (Droge & Feger, 1983; Feger & Bien, 1982) have proposed another method known as Ordinal Network Scaling (ONS), which represents rank orders of proximities by a network. All of these techniques hold the promise of providing a firmer theoretical and empirical foundation for network representations.

II. Pathfinder Networks

In an earlier paper (Schvaneveldt *et al.*, 1988), we presented a formal account of the graph-theoretic foundations of Pathfinder networks (PFNETs). Here we summarize these results along with a discussion of some of the properties of PFNETs. It is helpful to conceptualize proximity data as a complete network with the weight on each link equal to the proximity between the entities connected by link.[2] Call this network the DATANET. The DATANET is a direct representation of the proximities, but because of the density of links in the network, it is not very informative. An example of a DATANET is shown in Fig. 1A. The essential idea underlying Pathfinder networks is that a link in a DATANET is a link (with the same weight) in a PFNET if and only if the link is a minimum weight path in the DATANET. Equivalently, we can say that the PFNET has the same distance matrix as the DATANET, but the PFNET has the minimum number of links needed to yield that distance matrix.

A variety of different PFNETs can be derived from a given set of proximity data. A particular PFNET is determined by the values of two parameters, r and q. These two parameters represent generalizations of the usual definiticn of distances in networks. The r parameter determines how the weight of a path is computed from the weights on links in the path. The q parameter limits the number of links allowed in paths.

A. The r Parameter

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 weights of the links in the path. When link weights are obtained from empirical data, it may not be justifiable to compute path weight in this way because that computation assumes ratio scale measurement (Stevens, 1951). For computing distances in DATANETs, we need a distance function that will permit computations of distances in networks with different assumptions about the level of measurement associated with the proximities. From the perspective of deriving networks from proximities, such a distance function should preserve ordinal relationships between link weights and path weights for all permissible transformations of the proximities with different assumptions about the level of measurement associated with the proximities. Then, ordinal comparisons of path weights and link weights could be used to determine link membership in PFNETs.

A distance function with the required qualities can be defined by adapting the Minkowski distance measure to computing distances over paths

[2] The proximity estimates will define a complete network when the set of proximities is complete. Missing data can be handled by using infinity for missing values. Pairs of entities with infinite proximities will never be linked in any PFNET. This fact can also be used to prevent the linking of any two nodes simply by using infinite proximities for the appropriate pairs. PFNETs are not necessarily connected when some of the proximities are infinite.

A Proximity Data (Adjacency Matrix)

$$A = \begin{array}{cccccc} & A & B & C & D & E \\ A & 0 & 1 & 3 & 2 & 3 \\ B & 1 & 0 & 1 & 4 & 6 \\ C & 3 & 1 & 0 & 5 & 5 \\ D & 2 & 4 & 5 & 0 & 4 \\ E & 3 & 6 & 5 & 4 & 0 \end{array}$$

B Distance Matrix, r = ∞, q = 4

$$D^{\infty,4} = \begin{array}{cccccc} & A & B & C & D & E \\ A & 0 & 1 & 1 & 2 & 3 \\ B & 1 & 0 & 1 & 2 & 3 \\ C & 1 & 1 & 0 & 2 & 3 \\ D & 2 & 2 & 2 & 0 & 3 \\ E & 3 & 3 & 3 & 3 & 0 \end{array}$$

C Distance Matrix, r = 1, q = 4

$$D^{1,4} = \begin{array}{cccccc} & A & B & C & D & E \\ A & 0 & 1 & 2 & 2 & 3 \\ B & 1 & 0 & 1 & 3 & 4 \\ C & 2 & 1 & 0 & 4 & 5 \\ D & 2 & 3 & 4 & 0 & 4 \\ E & 3 & 4 & 5 & 4 & 0 \end{array}$$

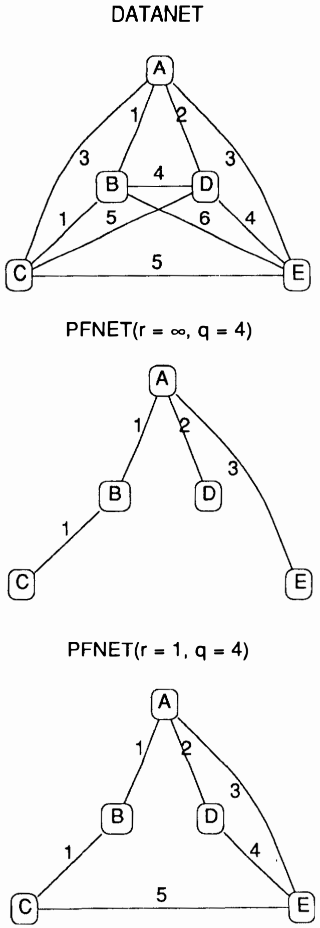

Fig. 1. Sample data and some networks derived by Pathfinder. A, A proximity matrix with (symmetrical) proximity estimates from the entity in the row to the entity in the column and the corresponding complete network. B, The *r* distance matrix using $r = \infty$ and $q = 4$ and the PFNET($r = \infty$, $q = 4$) for the data in Fig. 1A (also the minimal spanning tree for the complete network in Fig. 1A). C, The *r* distance matrix using $r = 1$ and $q = 4$ and the PFNET($r = 1$, $q = 4$) for the data in Fig. 1A.

in networks. It can easily be shown that the Minkowski *r* distance satisfies the requirements of a path algebra for networks as defined by Carre (1979). The *r* distance function replaces the normal sum with the *r* distance so that $x + y$ is replaced by $(x^r + y^r)^{1/r}$, $x \geqslant 0$, $y \geqslant 0$, $r \geqslant 1$. Given a path *P* consisting of *k* links with weights $w_1, w_2, \ldots, w_k$, the weight of path *P*, *w(P)* becomes:

$$w(P) = \left[\sum_{i=1}^{k} w_i^r\right]^{1/r} \qquad \text{where } r \geqslant 1, w_i \geqslant 0$$

Note that with $r = 1$, the function corresponds to simple addition (the usual definition of distances in networks). With $r = \infty$, the function is the maximum function. In fact,

$$\lim_{r \to \infty} [w_i^r + w_j^r]^{1/r} = \text{maximum}\,(w_{\mathrm{i}}, w_j)$$

Thus with $r = \infty$, computing network distances with the Minkowski r distance only requires maximum (as above) and minimum (for identifying geodesics or minimum weight paths) 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 link weights (proximities).

Another attractive property of the Minkowski r distance is that a single weight can be associated with a path regardless of segmentation. Given a set of path segments, S, which are mutually exclusive and exhaustive segments of path P (i.e., S is any decomposition of path P into subpaths.):

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

The use of the r parameter to compute path weights requires the assumption that the links 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, Dawes, & Tversky, 1970), r may be interpreted as a parameter of component weight. With $r = 1$, all components (links 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 (link) determines the weight of a path. The 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.

In summary, the r parameter for PFNETs is the value of r in the Minkowski r distance computation for the weight of a path as a function of the weights of links in the path. Variation of the r parameter can lead to different PFNETs, to which we return shortly.

B. The *q* Parameter

The distance matrix of a network is usually determined by finding the minimum weight paths regardless of the number of links in those paths. The *q* parameter is another generalization of this definition of network distance. This parameter places an upper limit on the number of links in paths used to determine the minimum distance between nodes in the DATANET. There are two reasons for using the *q* parameter, one psychological and the other representational. From a psychological perspective, there may be some limit on the number of links that could meaningfully connect nodes in a particular domain. This amounts to a limit in the chain of relations that can be constructed relating any two concepts in the domain. This limit can be incorporated into the network generation procedure with the *q* parameter. The representational motivation for the *q* parameter is that it provides a method for systematically controlling the density of links in PFNETs. Users of PFNETs may have various reasons for preferring networks of varying density. We examine this property in a following section. Thus, the *q* parameter further extends the family of PFNETs defined by Pathfinder. With *n* entities, possible values of *q* range over the integers from 1 to $n - 1$. With $q = 1$, the PFNET is the same as the DATANET, with $q = n - 1$, there is essentially no limit on the length (number of links) of paths because the longest possible path has $n - 1$ links.

C. Definition of Pathfinder Networks

With the two parameters *r* and *q*, a particular PFNET can be identified as PFNET*(r,q)*. We can now state the definition of Pathfinder networks precisely. Given a DATANET (proximities) with adjacency matrix $A = [a_{ij}]$ and a distance matrix $D^{r,q} = [d_{ij}]$ computed with parameters *r* and *q*: A link *(i,j)* in the DATANET is a link in the PFNET*(r,q)* if and only if

$$a_{ij} \neq \infty \qquad \text{and} \qquad d_{ij} = a_{ij},\ i \neq j$$

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

D. Some Properties of Pathfinder Networks

The *minimal PFNET* is PFNET($r = \infty$, $q = n - 1$). This PFNET has the fewest links of any PFNET for a particular set of data. With symmetrical proximity data (yielding undirected PFNETs), the edges in the minimal PFNET are the edges in the union of the edges in all minimal spanning trees (Kruskal, 1956; Dearholt, Schvaneveldt, & Durso, 1985) of the DATANET. The minimal PFNET will be the unique minimal spanning tree when there is such a unique tree. Certain patterns of ties in the proximity data may result in there being more than one tree, in which case the minimal PFNET will include all edges that are in any minimal spanning tree. Figure 1B shows the minimal PFNET for the proximity data in Fig. 1A. This PFNET is a tree (no cycles), and it is the minimal spanning tree for the complete network shown in Fig. 1A. There is also a close connection between minimal spanning trees and the single-link hierarchical clustering analysis (Johnson, 1967). The single-link clusters can be directly derived from the minimal PFNET using the link weights. However, it is not possible to recover the PFNET from the clustering solution because the details about which nodes are directly linked are not fully represented in the hierarchical clustering solution.

Using different r values to compute path weight will usually produce different PFNETs. For example, PFNET($r = 1, q = n - 1$) is the result of using the usual sum of the link weights in a path to define the path weight function. Figure 1C shows this PFNET for the proximity data in Fig. 1a. This PFNET has two additional links over the minimal PFNET, and the additional links necessarily introduce cycles.

The NETSCAL (Hutchinson, 1981) network generation method yields the same network as the PFNET($r = \infty$, $q = 2$). These PFNET parameters mean that only paths consisting of one or two links are examined in the DATANET when determining the minimum weight paths and, consequently, which links are to be included in the resulting network.

Decreasing either the r parameter or the q parameter leads to monotonic decreases in path weights and network distances. Because link membership in PFNETs is determined by the ordinal relationship of link weights and distances, decreasing either parameter can increase the number of links in a PFNET.

A network G' is *included in* in a network G if G and G' have the same nodes and the links in G' are a subset of the links in G. We also say that network *G includes* network G′. PFNET (r_1, q) is included in PFNET(r_2, q) if and only if $r_1 \geqslant r_2$. Similarly, PFNET(r, q_1) is included in PFNET(r, q_2) if and only if $q_1 \geqslant q_2$. The inclusion relationship means that the links in less dense networks are a subset of the links in more dense ones when the networks differ only in the value of one of the parameters. The links in PFNET($r = \infty$, $q = n - 1$) are found in all PFNETs.

E. Levels of Measurement

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

With measurement on a ratio scale (Stevens, 1951), the only allowable transformations that preserve the information in the scale values involve multiplication by a positive constant (i.e., a change of unit). Pathfinder networks will have the same structure (i.e., have exactly the same links) under multiplication of the proximity 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 links in the PFNET or other criteria.

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 (Stevens, 1951), Pathfinder will provide a unique PFNET structure only for $r = \infty$. That is, the same links will be present in the PFNET(∞, q) derived from any nondecreasing transformation of a particular set of proximities. Thus, the PFNET (∞, q) is a unique structure for levels of measurement ranging from ordinal through interval to ratio. It is the only unique structure with ordinal measurement.

It should be noted that transformations on proximities involving additive constants can lead to dramatic changes in the structure of PFNETs except for $r = \infty$.[3] Consequently, when using other values of r, it is particularly important for the proximity estimates to be measured on a scale with a "true" zero.

F. Distances in Pathfinder Networks

Once a PFNET has been obtained, it is often of interest to derive measures of distance between nodes in the network. For example, these dis-

[3] An additive constant has this effect because it is included in the weight on each link in a path. When these weights are summed, the constant is included as many times as there are links in a path. Because paths have varying numbers of links, the constant has a variable

tances can be used to predict performance on tasks involving the concepts corresponding to the nodes or to determine the fit between the network distances and the proximities. However, the scale of measurement underlying link weights places constraints on computing distances in PFNETs just as it does for computing distances in DATANETs. With ratio scale measurement, there are several options for determining distances in PFNETs including using the usual sum of the link weights. With ordinal scale measurement, the options are more limited. Here we focus on the more difficult ordinal measurement case. Methods of computing distance when the proximities are measured on an ordinal scale should yield invariant measures of distances with any monotonic transformation of the proximities. As discussed above, only PFNETs computed with $r = \infty$ are appropriate with ordinal data.

One method we have found useful involves concentrating on the structure of the PFNET by treating the network as a graph. This approach requires ignoring link weights, or, equivalently, giving each link a weight of one (1). With this method, the distance between two nodes is the (minimum) number of links connecting the nodes. Importantly, these distances will be the same whenever the same links are present in the PFNET. Another approach to using only ordinal information is to rank-order the link weights and compute distances using ranks. These ranks would be preserved for any monotonic transformation of the proximity data and, consequently, so would distances computed with ranks.

Once we have distances from PFNETs, we can determine the fit between these distances and the original proximities by computing the correlation between them. If ordinal measurement is involved, rank-order correlations should be used. The fit between PFNET distances and the proximity data provides one method for selecting one of the possible PFNETs. By simultaneously considering the fit of the PFNET to the data and the density (number of links) of the PFNET, it is possible to choose a PFNET that is optimal in the sense of maximizing fit while minimizing density. This is similar to the elbow criterion used in MDS to pick the appropriate dimensionality. In both cases, the goal is to account for a maximum of variability in the data with a minimum number of parameters. Such statistical determination of the correct solution should be only one criterion used by the researcher. With Pathfinder, as well as other methods, it is the interpretability of the solution that is, after all, the pri-

effect. Of course this problem does not occur when the path weight is determined by taking the maximum of the weights of links in the path, that is, when $r = \infty$.

mary goal. In some of our own work, we have found that the network that best fits the proximity data is not always the one that produces the best results using some other criterion external to the proximity data.

G. Pathfinder Algorithms

We have implemented various algorithms for deriving PFNETs in several computer languages running on several different computers.[4] The derivation of Pathfinder networks requires computing the distance matrix of a complete (or nearly complete) network (see Aho, Hopcroft, & Ullman, 1974). With n nodes (or entities), the best general algorithm we have implemented to date has time complexity of $O(n^3 \log q)$. The best special case algorithm has time complexity of $O(n^3)$. Although complexity at these levels is prohibitive for rapid computation on large networks, it is quite manageable for occasional derivations of networks with hundreds of nodes. On a few occasions, we have derived networks with over 2000 nodes. Several potential applications of Pathfinder require analysis of problems of this size or smaller. Many of our studies have been conducted with networks consisting of 30 or fewer nodes.

III. Applications of Pathfinder Networks

We have investigated Pathfinder network structures in a variety of domains. The examples presented here were selected to illustrate the results we have obtained using Pathfinder and to highlight some of the unique properties of the networks. The examples include demonstrations, confirmations of theoretical analyses, and validation tests of predictions made using the network structures.

The methods used to obtain the proximity data used in the Pathfinder analyses are straightforward and analogous to methods employed to obtain data for MDS or cluster analysis. Except in examples using data borrowed from the literature, each proximity matrix submitted to Pathfinder represented the mean judgments of a number of subjects asked to judge the similarity or relatedness of all pairwise combinations of stimuli using a scale ranging from 0 to 9. Stimulus presentation was randomized and controlled by a microcomputer, which also recorded the subjects' judgments.

[4] Programs have been written in Pascal, C, LISP, and APL. Various versions of the programs run on IBM PC, Apple Macintosh, and SUN Microsystems. Information on obtaining programs is available from Interlink, Inc., P.O. Box 4086 UPB, Las Cruces, NM 88003-4086.

A. Natural Concepts

All pairwise combinations of 25 natural concepts were rated for degree of relatedness by 24 students in introductory psychology courses. The concepts and the average (multiplied by 10) pairwise proximities are shown in Fig. 2. Note that the proximities are symmetric, that is, the proximity of concept *i* and concept *j* is the same as the proximity of *j* and *i* for all *i* and *j*. With symmetric data, Pathfinder produces undirected networks. The concepts were chosen to represent a variety of relationships including categories, properties, habitats, and similarities.

There is a family of Pathfinder networks for any set of data. The particular network selected from this family will depend on assumptions about the empirical data and on decisions about the number of links (*q*) permitted in paths considered in the DATANET in finding minimum distances.

Concepts

A.	living thing	F.	robin	K.	deer	P.	plant	U.	rose
B.	animal	G.	chicken	L.	bat	Q.	leaves	V.	daisy
C.	blood	H.	mammal	M.	antlers	R.	tree	W.	color
D.	bird	I.	hair	N.	hooves	S.	cottonwood	X.	green
E.	feathers	J.	dog	O.	frog	T.	flower	Y.	red

Proximities

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
A	0	13	29	18	51	23	17	18	45	15	15	22	41	48	23	20	33	18	28	22	30	31	45	35	62
B	13	0	26	25	44	34	23	15	33	13	11	28	31	35	28	47	65	49	58	49	64	55	53	69	63
C	29	26	0	47	54	43	43	27	55	39	37	35	48	54	51	73	64	74	76	70	47	78	35	76	15
D	18	25	47	0	8	12	18	36	65	35	41	22	73	72	48	48	53	26	46	50	48	54	45	65	49
E	51	44	54	8	0	17	15	62	36	73	73	70	56	56	72	73	53	67	62	62	73	68	45	59	43
F	23	34	43	12	17	0	27	47	67	44	42	32	65	74	48	55	55	26	37	47	49	53	39	74	27
G	17	23	43	18	15	27	0	47	67	41	41	45	68	76	47	63	72	58	57	60	63	67	55	78	56
H	18	15	27	36	62	47	47	0	33	20	24	24	35	35	49	54	59	54	60	60	63	63	56	72	62
I	45	33	55	65	36	67	67	33	0	20	44	47	49	53	80	64	55	71	69	73	75	72	49	75	58
J	15	13	39	35	73	44	41	20	20	0	35	43	66	61	42	57	68	46	57	63	64	57	51	77	62
K	15	11	37	41	73	42	41	24	44	35	0	46	11	21	44	53	59	50	52	55	58	64	50	73	69
L	22	28	35	22	70	32	45	24	47	43	46	0	63	69	43	64	66	43	58	62	70	66	66	75	70
M	41	31	48	73	56	65	68	35	49	66	11	63	0	27	71	67	67	58	61	64	71	75	64	74	69
N	48	35	54	72	56	74	76	35	53	61	21	69	27	0	73	73	66	75	74	75	74	74	64	78	73
O	23	28	51	48	72	48	47	49	80	42	44	43	71	73	0	52	64	51	58	62	63	63	50	24	78
P	20	47	73	48	73	55	63	54	64	57	53	64	67	73	52	0	16	13	26	9	17	18	40	11	47
Q	33	65	64	53	53	55	72	59	55	68	59	66	67	66	64	16	0	12	27	23	32	34	38	16	44
R	18	49	74	26	67	26	58	54	71	46	50	43	58	75	51	13	12	0	11	29	35	34	52	17	61
S	28	58	76	46	62	37	57	60	69	57	52	58	61	74	58	26	27	11	0	32	40	39	56	45	65
T	22	49	70	50	62	47	60	60	73	63	55	62	64	75	62	9	23	29	32	0	8	9	27	33	30
U	30	64	47	48	73	49	63	63	75	64	58	70	71	74	63	17	32	35	40	8	0	19	23	60	14
V	31	55	78	54	68	53	67	63	72	57	64	66	75	74	63	18	34	34	39	9	19	0	43	49	58
W	45	53	35	45	45	39	55	56	49	51	50	66	64	64	50	40	38	52	56	27	23	43	0	10	11
X	35	69	76	65	59	74	78	72	75	77	73	75	74	78	24	11	16	17	45	33	60	49	10	0	32
Y	62	63	15	49	43	27	56	62	58	62	69	70	69	73	78	47	44	61	65	30	14	58	11	32	0

Fig. 2. Average pairwise proximity estimates for 25 natural concepts.

TABLE I

NUMBER OF LINKS IN PATHFINDER NETWORKS OF NATURAL CONCEPTS AS A FUNCTION OF *r* AND *q*

	q			
r	2	3	4	24
1	119	104	103	103
1.01	102	89	87	87
1.05	95	83	81	81
1.1	86	75	70	70
1.15	76	66	62	61
1.2	72	65	60	59
1.4	63	53	53	52
1.6	56	51	51	50
1.8	50	47	45	45
2	47	44	42	42
3	39	37	36	34
4	35	31	31	29
5	32	30	29	27
6	32	30	28	26
∞	32	28	27	25

These factors have a direct and predictable influence on the density of the network. Table I presents the number of links in each network as a function of the values of the *r* and *q* parameters.[5]

The maximum density occurs when path weights are computed by summing link weights and only paths of two links or less are considered in finding minimum-length paths, that is, PFNET($r = 1$, $q = 2$). The minimum density results from using the maximum link weight in a path to determine the weight of a path and paths of any number of links are examined, that is, PFNET($r = \infty$, $q = 24$). Table I shows a clear relationship between these two parameters and the resulting density of the network; density is weakly monotonic with *r* and *q*. In addition, the family maintains qualitative relations among its members. Links in the less dense members will also be found in the more dense ones.

[5] The proximity data collected for the natural concepts would only justify the use of $r = \infty$ in deriving PFNETs. Other values of *r* are used only to illustrate the systematic variation in density with a particular set of data. For detailed analyses of these networks, we shall confine our discussion to PFNETs with $r = \infty$.

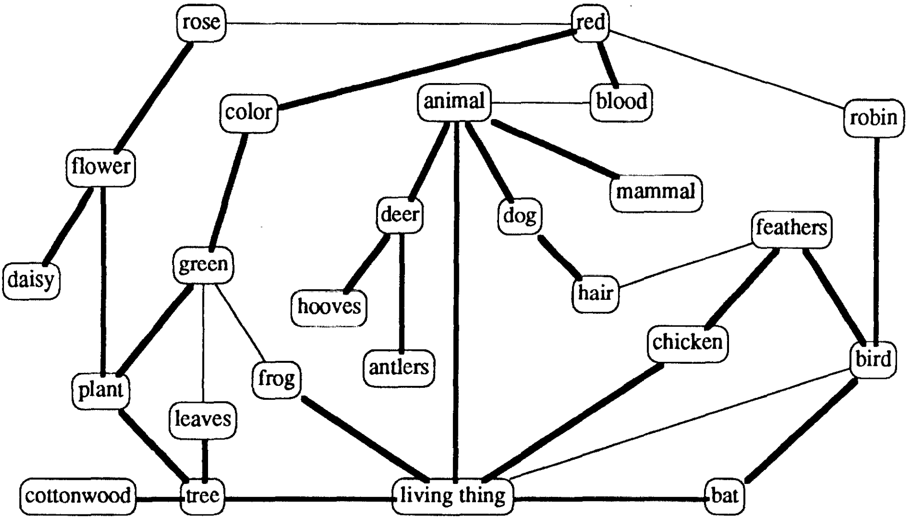

Fig. 3. Networks derived by Pathfinder from the data shown in Fig. 2 using $r = \infty$. The heavy links are from PFNET($r = \infty$, $q = n - 1$). The thin links are added in the PFNET ($r = \infty$, $q = 2$) solution.

Figure 3 displays two networks computed with $r = \infty$. The least dense network shows a number of interesting connections. The PFNET($r = \infty$, $q = 24$) for these data yielded 25 links compared to the minimum $n - 1 = 24$ links (a tree). The additional link assures the presence of one cycle and hence the network is not a tree. The cycle is *living thing–bat–bird–feathers–chicken–living thing.* The cycle occurs because of the tie in the data for *living thing–bat* and *bat–bird.* Both of these links are included to insure that the resulting structure is unique. A minimal spanning tree would result from removing either of these links.[6] Several types of relationships appear to be represented in this network. *Bird,* for example, connects to both the concept *robin* and the property *feathers,* suggesting the links might be labeled *isa* and *has,* respectively. The most general concept, *living thing,* is involved in several connections: in graph-theoretic terms it has a degree of 4. The closest node pairs are *bird–feathers* and *flower–rose.* The longest link is *living thing–frog.*

Category members that one might view as typical of a superordinate

[6] Link weights are omitted from most of our figures to enhance their appearance. In this case, the link weights can be obtained from Fig. 2 by using the proximities for the appropriate pairs of concepts.

category tend to be linked to that category directly, whereas atypical members tend to be connected via a path of concepts. For example, *robin* links directly to *bird,* whereas the path *chicken–feathers–bird* connects *chicken* with *bird.* Similarly, the typical animals (i.e., *dog* and *deer*) have direct links to *animal* while the less typical ones have multiple-link paths, usually through *living thing*. Along the same lines, the scientific category *mammal* and its members are always connected through a path and not directly linked. Even in networks of higher density, only *bat* links to *mammal.* Perhaps this link represents a connection established in school to prevent the inference that a bat is a bird.

As we increase the density of the networks by decreasing q, we see that the links added to the network continue to suggest readily interpretable relations. PFNET($r = \infty$, $q = 2$) adds *green–frog, green–leaves, red–rose, red–robin, animal–blood,* and *feathers–hair.*

By way of comparison, the best MDS solution was a three-dimensional space. Optimal dimensionality was determined by a number of factors: stress and R^2 tended to elbow at two or three dimensions; the addition of a third dimension clarified the prior ones and was itself interpretable (Shepard, 1974); and the Isaac and Poor (1974) procedure suggested three dimensions. The dimensions appear to be plant–animal, entities–properties, and hueless–colorful. This type of global information could not be extracted easily from the network solutions. However, comparison of MDS and Pathfinder at a more local level suggests that Pathfinder has more accurately captured the pairwise relations.

For example, in attempting to satisfy all of the constraints in the proximities, MDS positioned the concept *chicken* far from the property *feathers,* but the two are linked in even the least dense network. In contrast, the concept *chicken* is close to the concept *bat* although they are not linked in even the most dense network. The network appears to agree better with intuition and with the mean proximities from our subjects: The *chicken–feathers* pair was very close (15) compared with *chicken–bat* (45).

B. Experts and Novices

Several studies attest to differences in knowledge organization in experts and novices (Chase & Simon, 1973; Chi, Feltovich, & Glaser, 1981; McKeithen, Reitman, Rueter, & Hirtle, 1981; Reitman, 1976; Schvaneveldt *et al.,* 1985). Can Pathfinder capture these expert–novice differences in conceptual structure? As a first step in answering this question, we obtained judgments of relatedness for all pairs of the concepts in Fig. 2 from 12 graduate students in biology at New Mexico State University.

From the average judgments, we derived Pathfinder networks. These networks can be compared with the undergraduate psychology student networks for the same concepts.

To compare the two groups, a PFNET was selected for each group using the fit (rank-order correlation of the proximities and the minimum number of links between nodes in a PFNET) for a number of PFNETs generated with $r = \infty$. These correlations are shown in Fig. 4 as a function of the density (number of links) of the PFNETs.

There are apparent elbows in the functions. Below the elbows, increases in fit can be obtained with small increases in density. Above the elbows, much larger increases in density are required for comparable increases in fit. The elbows occur with the PFNET($r = \infty$, $q = 3$) for the students and PFNET($r = \infty$, $q = 5$) for the biologists. The selected networks are shown in Fig. 5.

There are several similarities and differences in the networks for the undergraduate psychology students and the graduate biology students. The undergraduate network has 28 links, and the graduate network has

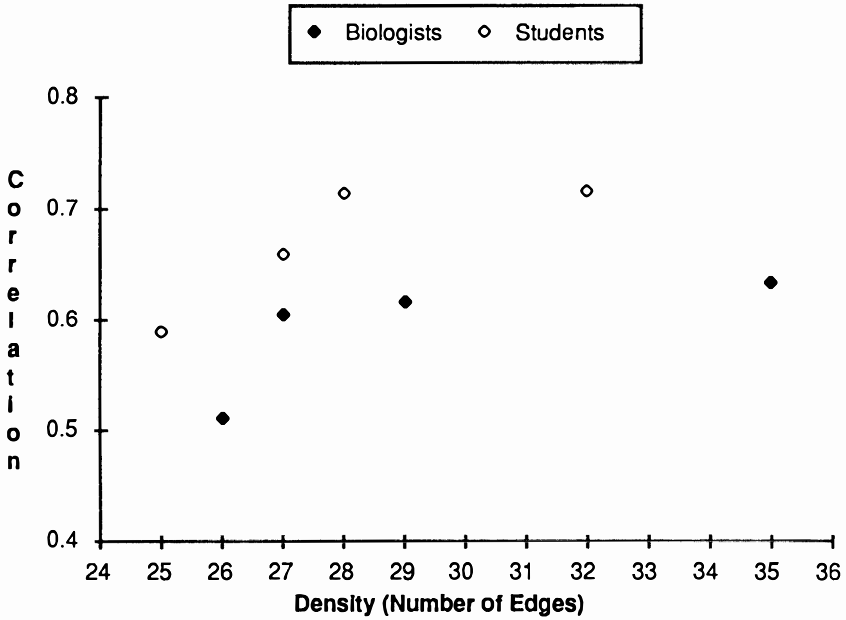

Fig. 4. Fit (rank-order correlation) between the student and biologist proximity data and distances (number of links between nodes) derived from various Pathfinder networks (varying values of q with $r = \infty$) as a function of network density.

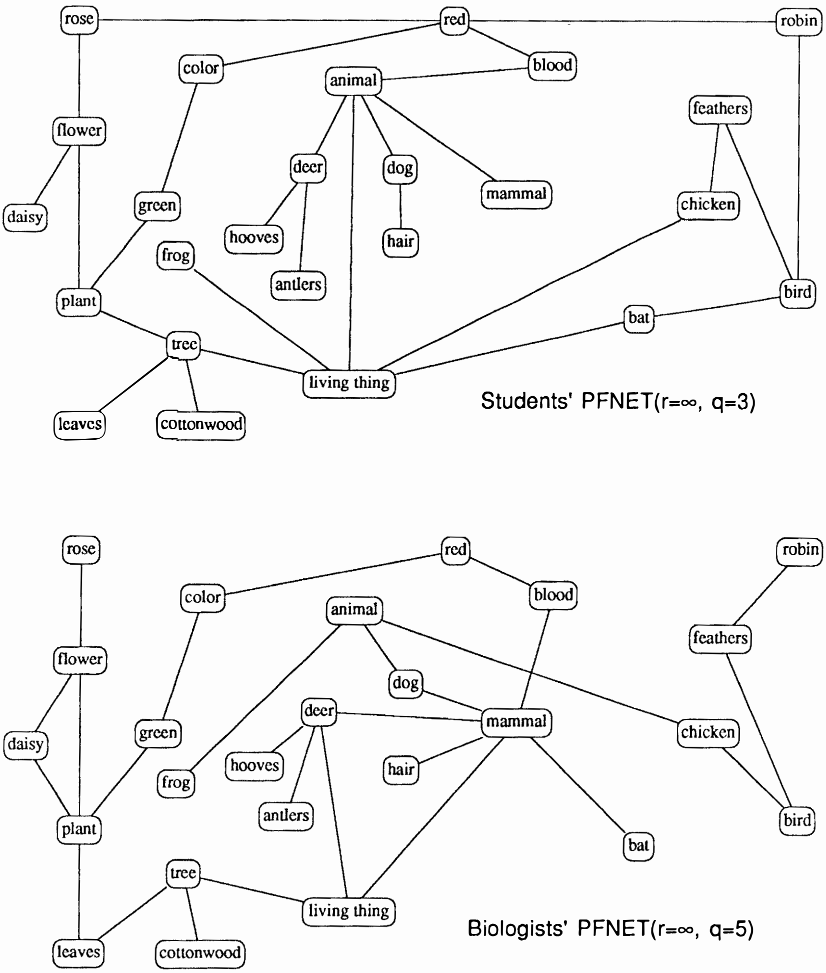

Fig. 5. The "best fitting with minimum density" or "elbow" networks from Fig. 4 for students [PFNET($r = \infty$, $q = 3$)] and biologists [PFNET($r = \infty$, $q = 5$)] for the natural kind concepts shown in Fig. 2.

27. The two networks share 14 links. We have found in several informal tests that people can quite easily associate these networks with the appropriate groups. Perhaps the most diagnostic difference can be found in the role played by *mammal* in the two networks. For the undergraduates, *mammal* is only connected to *animal* while the graduate biology students have *mammal* connected to *deer, dog, hair, bat,* and *blood.* Not surpris-

ingly, *mammal* is a much richer concept for the biologists. The Pathfinder networks help to highlight these conceptual differences between experts and novices.

C. Basic Level Categories

Rosch's work on basic level categories represents an important contribution to our understanding of category structure. Rather than assuming that category structure follows a strictly hierarchical superordinate–subordinate structure, Rosch has postulated that there exists a psychologically special level of categorization. She has supplied a wealth of empirical evidence supporting this view of basic level categories.

Rosch (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) has argued that "the basic level of abstraction in a taxonomy is the level at which categories carry the most information, possess the highest cue validity and are thus, the most differentiated from one another." At the basic level, objects share a maximum number of attributes while sharing a minimum number of attributes with objects in contrasting categories. One characteristic of basic level categories, which is apparent without recourse to the empirical work, is that the basic level term tends to be applied in identifying an object. For example, unless a request to identify an object implies the desire for detailed information, a person will call a chair a *chair* and not *furniture* or *wicker chair.*

In particular, Rosch showed that the categories *bird, fish,* and *tree* exhibited the properties of basic level categories, whereas *musical instruments, clothing,* and *fruit* (among others) had the properties of superordinate categories. In this section we discuss the application of Pathfinder to these six categories and four additional ones.

In addition, this section highlights the ability of Pathfinder to accommodate asymmetrical proximities. Several of the criticisms of various scaling procedures stem from their inability to represent asymmetrical relations. Tversky (1977), for example, contends that such asymmetries are not simply perturbations in data but that they have meaningful psychological interpretations. Certainly, the logical relationships between a category and its members are asymmetric. Such asymmetries are also apparent in association norms.

We began with the Marshall and Cofer (1970) report of the Connecticut norms (Cohen, Bousfield, & Whitmarsh, 1957). These norms are controlled four-response associates to category labels from 400 individuals. Responses of nonzero frequency for our 10 categories were noted. We then searched the Marshall and Cofer (1970) single-response free association norms of 100 people for word associations to these responses. Any response to the Connecticut norms that appeared in the Marshall and Cofer norms was retained. In this way we obtained n stimuli that included the category name and any category members that occurred as responses to the category name and were also used as stimuli in free association. We then created $n \times n$ matrices for each category where the cell was the proportion of people giving a response for a stimulus subtracted from 1.0.

The resulting matrices are clearly asymmetrical. For example, *thrush* was given as a response to *bird* on only 3 of 1600 opportunities, whereas *bird* represented 31% of the responses to *thrush*. In addition, some responses were never given to some stimuli, yielding infinite proximities for these cells. Pathfinder handles an infinite proximity between two concepts by not permitting a link between the concepts. In principle then, Pathfinder is able to construct disconnected networks. We discovered in our attempts to apply ordinal MDS to these data that the algorithm had a number of difficulties, perhaps because of the infinite values which were treated as missing data. Thus, two-dimensional MDS solutions were as appropriate as higher-dimensional solutions, but none of the solutions were very good. Although we compare the MDS and Pathfinder solutions, one may choose instead to assume that MDS cannot supply a reasonable fit for these data.

The PFNETs for the six categories investigated by Rosch appear in Fig. 6. These networks are directed PFNET($r = \infty$, $q = n - 1$) networks.

The MDS solutions (not shown) did capture some of the category structure. These solutions placed the category label in the center of the space and surrounded the label with the instances. However, it tended to do this for both the superordinate and the basic level categories. Pathfinder, on the other hand, tends to show a star-shaped network for the basic level categories. This star-shaped network is less apparent in the superordinate cases.

We can quantify the *starness* in each pattern by calculating the relative degree of the category node. Because the network is directed, each node has both an in-degree and an out-degree. The *in-degree* is the number of directed links terminating on the node, and the *out-degree* is the number of directed links initiating from the node. The sum of these two is the total degree of the node. Dividing the total degree of the category node by the total number of links in the network gives us the percentage of links that connect with the category node (relative degree), which we used as an index of the starness.

The starness indices are presented in Table II for the six categories used by Rosch and for four additional categories she did not consider. The basic level categories show greater involvement of the category label in the network. Based on the starness indices derived in the same way

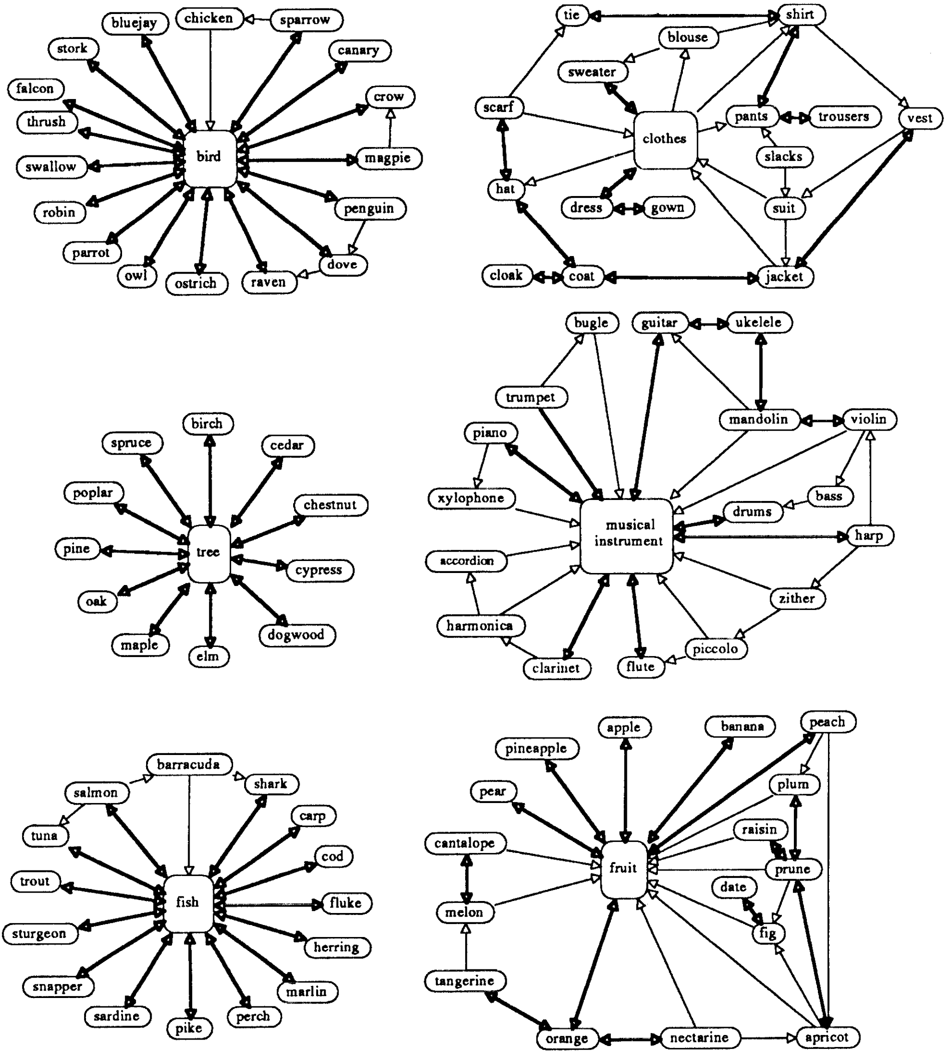

Fig. 6. Directed Pathfinder networks for various categories and category members.

for the categories *flower, profession,* and *body (parts),* we would classify them as basic level, superordinate, and superordinate terms, respectively. With a starness index of .70, the appropriate classification for the category *metal* is uncertain. Its starness falls squarely in the middle of the values for basic level and superordinate categories.

TABLE II

STARNESS OF VARIOUS CATEGORY NETWORKS

Category	Type	Starness[a]
Fish	Basic level[b,c]	.90
Bird	Basic level[b,c]	.89
Tree	Basic level[b,c]	1.00
Musical instrument	Superordinate[b,c]	.56
Fruit	Superordinate[b,c]	.50
Clothes	Superordinate[b,c]	.31
Flower	Basic level[c]	1.00
Profession	Superordinate[c]	.50
Body (parts)	Superordinate[c]	.37
Metal	?[d]	.70

[a] Starness is the proportion of links in the network directly connecting with the category name node.
[b] Rosch's classification.
[c] Starness classification.
[d] Uncertain classification.

The Pathfinder networks revealed structural differences among categories that have been shown to have different characteristic properties and that yield different results in a variety of psychological experiments. Although MDS captured some of the category information by placing the category concept in the center of the space, it did not uncover differences between superordinate and basic level concepts. Pathfinder yielded networks for basic level categories in which the category concept had a high total degree, in some cases accounting for 100% of the links in the network. Networks of superordinate categories, on the other hand, yielded category concepts with relatively lower total degree.

D. THE COLOR CIRCLE AND THE COLOR CYCLE

The classes of concepts considered thus far clearly refer to discrete entities. They are also complex in the sense that one would have expected Pathfinder networks with a number of connections if in fact it did capture part of the latent structure in subjects' similarity ratings or word associations.

The next set of data was collected by Ekman (1954) in a study of color perception. The data were borrowed by Shepard (1962b) in his development of nonmetric MDS, in which case the data yielded a two-dimensional color circle. Figure 7 presents the PFNET resulting from Pathfinder superimposed on the two-dimensional Shepard solution.

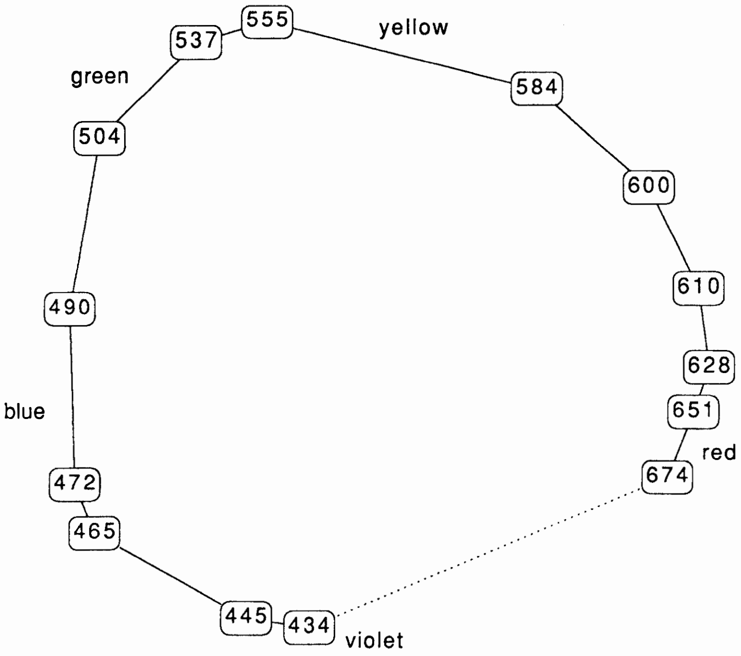

Fig. 7. Pathfinder network and a two-dimensional MDS for judgments of similarity of colors. Wavelengths ranging from 434 to 674 mμ were judged. The PFNET($r = \infty$, $q = n - 1$) is shown with solid lines. The dotted line is added in the PFNET($r = \infty$, $q = 2$) solution.

The solid lines represent the PFNET($r = \infty$, $q = n - 1$) network (the minimal spanning tree) for the data, and the dotted line is the only link added to create the network, PFNET($r = \infty$, $q = 2$). The tree captures the psychological judgments that have a monotonic relation to physical wavelength; the PFNET($r = \infty$, $q = 2$) adds a single link that highlights the psychological similarity of two physically very different wavelengths. Shepard added exactly these lines to his MDS solution in order to highlight the circular nature of his solution. Pathfinder produces, algorithmically, the same lines as Shepard added to his MDS solution. In this case, a single cycle in the PFNET corresponds to a circle in space.

E. Unidimensional Network

We turn next to a coherent set of concepts that seemed neither complex nor discrete. We wondered what Pathfinder networks would reveal for a set of concepts that had a clear underlying dimension. We chose a set of

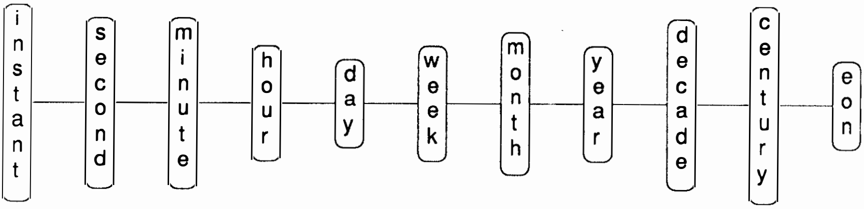

Fig. 8. PFNET($r = \infty$, $q = n - 1$) for judged similarity of temporal concepts.

words that varied on a dimension signifying time, or more accurately length of time. The set of 11 concepts ranging from *instant* to *eon* appears in Fig. 8. The data were obtained by averaging the pairwise relatedness judgments of 24 subjects. The network is a PFNET($r = \infty$, $q = 10$) solution.

Pathfinder produced a pattern at the opposite extreme from the star pattern we observed for basic level categories. Rather, a single path *(instant–second–minute–hour–day–week–month–year–decade–century–eon)* that perfectly mirrors the logical relations among the concepts was obtained. In more dense graphs, such long paths may suggest some underlying dimension, which may lead the researcher to more spatial algorithms in order to ascertain the nature of the dimension. However, as we argued earlier, the spatial algorithms will tend to distort some relations in order to find the best fit to the data. For the time concepts, a one-dimensional MDS solution did not reproduce the logical string of concepts that Pathfinder produced. The MDS solution placed the concept *eon* before the concept *century* and after the concept *decade*. Apparently, the constraints from all of the pairs that influence the MDS solution were sufficient to alter the order of two of the items. Pathfinder, with its emphasis on the smaller proximities, preserved this ordering, which was inherent in the pairwise data.

F. Network of a Script

Among the knowledge representations of current interest in cognitive science are schema or frame structures (Minsky, 1975; Rumelhart & Ortony, 1977). Scripts (Schank & Abelson, 1977) are one type of schemata that pertain to knowledge about recurring activities. It is of theoretical interest to examine how scripts and network structures relate. Scripts are also of interest here because of their property of having a number of complex relations in addition to an assumed underlying temporal dimension.

The concepts composing the restaurant script were taken from a disser-

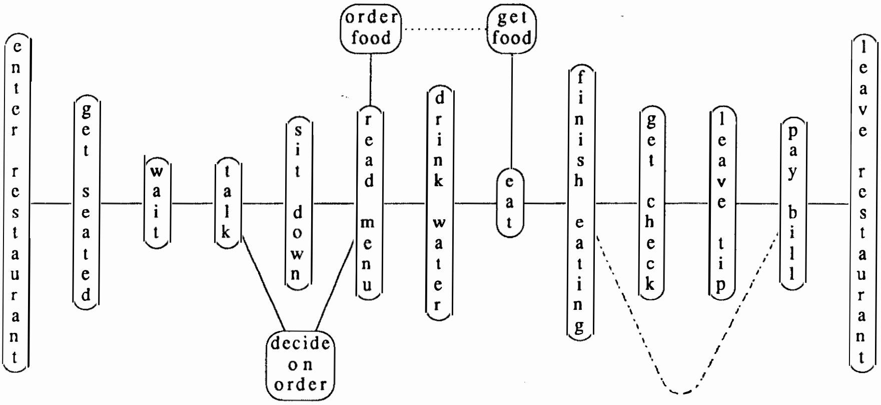

Fig. 9. Network for the judged relatedness of activities in restaurants (restaurant script). The solid links are from PFNET($r = \infty$, $q = n - 1$). The dotted links are added in the PFNET($r = \infty$, $q = 2$) solution.

tation by Maxwell (1983). Maxwell had undergraduate psychology majors generate a sequence of ordered actions that describe what people usually do when they go to a restaurant. We selected the 16 most frequent actions and had subjects judge the relatedness of the members of all possible pairs of these actions. We then obtained a PFNET($r = \infty$, $q = 15$) solution from Pathfinder. Figure 9 shows the network; the solid lines are links from the PFNET($r = \infty$, $q = 15$) solution.

The temporal dimension seems to have been revealed by the PFNET($r = \infty$, $q = 15$) solution. The events in going to a restaurant proceed in a reasonably linear fashion from *enter restaurant* to *leave restaurant*. There is one cycle in the PFNET($r = \infty$, $q = 15$) network that may be interpreted as cotemporal behaviors, alternate paths, or a point in going to restaurants at which possibilities vary.

Additional variations were found in the more dense PFNET($r = \infty$, $q = 2$) solution. The dotted lines in Fig. 9 are the links added by PFNET($r = \infty$, $q = 2$). The PFNET($r = \infty$, $q = 2$) solution added a link between the two actions, *order food* and *get food,* creating another cycle. It also added another cycle by connecting *finish eating* and *pay bill.*

The added density with variation in the q parameter raises the general problem of selecting among several network solutions. As a descriptive tool, Pathfinder can provide several ways of looking at the data and, of course, such exploration is entirely consistent with the goals of description. In our experience, it has often been helpful to begin with the least dense network and to consider additional links after the core links have been analyzed. The rules on when to stop, however, are not easy to define for all cases. One can use criteria determined by (1) the nature of measurement in the data (ordinal data require $r = \infty$); (2) the interpretability of the links; and (3) the function relating fit to density. Each of these criteria may have its place in the selection of particular networks.

The networks for the restaurant script revealed a strong underlying temporal dimension for the concepts. The more dense network also produced various cycles within the script. Using the network as a guide for going to a restaurant would not lead one far astray. The network provides some relatively invariant sequences of behavior and three points at which more than one behavior is appropriate.

G. Classifying Individuals

In this section we review work showing that Pathfinder supplies information about the cognitive structure of individuals that is useful in classifying them into their appropriate groups. In particular, this work showed that expert and novice fighter pilots could be classified on the basis of individuals' networks of flight-related concepts.

Schvaneveldt *et al.* (1985) asked expert (USAF instructor pilots and Air National Guard pilots) and novice (USAF undergraduate pilot trainees) fighter pilots to judge the relatedness of concepts taken from two domains: an air–air combat scenario (split-plane maneuvers) and an air–ground combat scenario (strafe run). Schvaneveldt *et al.* reported a number of uses of Pathfinder networks, but what is of interest here is their use of Pathfinder to classify an individual as an expert or novice.

A PFNET($r = \infty$, $q = n - 1$) was computed for each individual for each scenario. A vector was then created for each of these networks. This vector consisted of a series of zeros and ones for all possible pairs of concepts. A zero signified that the pair was not linked in the network, and a one signified that there was a link for that concept pair for that subject. These vectors were then used in a pattern classification procedure (Nilsson, 1965) of the type used by artificial intelligence devices to segment patterns.

A pattern classification system was defined using all but one expert and all but one novice. The classification system then attempted to classify the remaining two unknown individuals. This procedure was repeated a number of times by making certain that classification was attempted for each possible pair of unknown individuals. The percentage of correct classifications can then be computed and used as an index of the success of the vectors at capturing differences among the groups.

Network vectors were created as described above and vectors were also created based on the original ratings of the subjects and on the distances between concepts in MDS solutions. The ratings vectors simply consisted of the rating given by each individual to each pair of the 30 concepts. The MDS vector consisted of the Euclidean distances between concepts for all pairs in the MDS solution for each subject. Figure 10 shows the percentage correct classification for each type of vector for various pairs of groups and type of maneuver. Classification based on the network or on MDS was superior to classification based on the original ratings in each case, suggesting that both Pathfinder and MDS were successful at revealing the latent structure in the relatedness ratings that allows for a distinction among groups. In addition, MDS was better than Pathfinder in two of the four comparisons, equal in one, and inferior in one. Thus, MDS and Pathfinder both captured important structural differences, but the MDS distances led to somewhat more success.

These classification experiments show that the Pathfinder and MDS scaling techniques both extract information characteristic of expertise that is not directly available in the original ratings. We take this as a form of validation for both of these procedures.

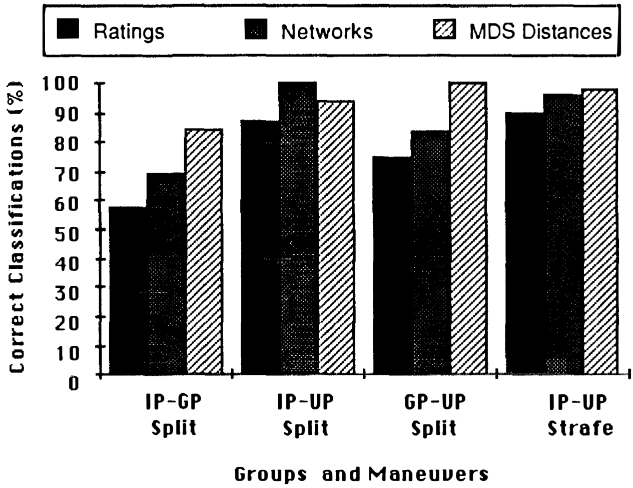

Fig. 10. Pattern classification analysis for identifying pilots using various measures. IP, instructors; GP, Air National Guard pilots; UP, Pilot trainees; Split, split-plane maneuvers; Strafe, low angle strafe maneuver.

H. Recall Studies

The value of the information extracted by Pathfinder has proven useful in some psychologically meaningful ways. One such investigation of Pathfinder made use of the fact that recall benefits from organization (e.g., Bousfield, 1953; Bower, 1972). In one study, we (Cooke, Durso, & Schvaneveldt, 1986) created organized and unorganized lists from the set of natural concepts discussed earlier (see Fig. 2). Organized lists were defined either as lists containing pairs of adjacent words linked in the network but not close in MDS space (network list) or as lists containing pairs of adjacent words close in MDS space but not linked in the network (MDS list). Control lists were created from these by scrambling the respective lists so that adjacent pairs were neither linked nor close.

Subjects were asked to learn the lists, and the number of trials to one perfect serial recall was recorded. As expected from other work, the organized lists were learned more quickly than the controls. However, there was also an effect of representation: The network lists were learned more rapidly than the MDS lists. The network advantage was present for the organized lists, but not the control lists, suggesting that it was not the words that facilitated learning but rather the organization of the words in the lists. In a replication, Cooke *et al.* created an MDS list and a network list for the same set of words, and, again, the network-organized list yielded more rapid learning.

Finally, 13 of the words were presented to 60 subjects for free recall. The order of the words was randomized for each trial for each subject, and trials were continued until all of the words were correctly recalled on one trial. All of the pairwise distances (number of intervening words) between words in the final recall order were determined for each subject, and these distances were averaged across subjects. These distances were then correlated with the earlier ratings of all of the concept pairs (also in Fig. 2). The recall distances were also correlated with the distances extracted from various MDS representations of the items and with the distances extracted from various Pathfinder networks. These scaling solutions were derived from the original rating data.

The average correlations between recall distances and the other measures were .56, .44, and .55, for ratings, MDS distances, and network distances, respectively. Perhaps of more interest were the partial correlations. The average correlation between recall distances and network distances was .34 with the ratings partialed out. The correlation between recall distances and MDS distances was −.004 with the ratings partialed out.

The partial correlations are particularly revealing because the network structure and the associated distances are derived from the ratings. The correlation of recall distances and network distances independently of the original rating data suggests that Pathfinder extracts important structural information from the rating data. We suspect that the gain from the Pathfinder method is due to the emphasis on closely related items in determining the network structure. The distances between more remotely related items are then derived from combinations of distances between linked items. People may be better at estimating the psychological distance of closely related items than they are with more distantly related items.

Results of the Cooke *et al.* study demonstrate that the information extracted by Pathfinder is useful in predicting recall orders and in generating easy-to-learn lists. Apparently the types of relations utilized by individuals when they attempt to remember a series of events is part of the information revealed by Pathfinder networks.

IV. Discussion and Future Lines of Investigation

Networks have several properties that should be of value in representing the structure in proximity data. Networks reduce a large number of pairwise proximities to a smaller set of links. Understanding of the data is simplified by this reduction. Networks highlight the local relationships among the entities represented. They are also capable of revealing several particular structures such as trees (including hierarchical structures, stars, and linear paths), cliques (a completely connected subgraph), and cycles.

Compared to spatial scaling methods, networks focus on the closely related (low dissimilarity or high similarity) entities. As a result, the pairwise information may be better represented than it is in spatial methods such as MDS. In contrast, spatial methods are probably superior in extracting global properties of a set of entities in the form of dimensions of the space. In some cases, the pairwise relations are distorted by MDS as all constraints in the pairwise data contribute to the location of entities in the space. Based as it is on finding minimum weight paths connecting entities, Pathfinder tends to give greater weight to the smaller values in the proximities.

Other nonspatial scaling methods such as hierarchical cluster analysis (Johnson, 1967), weighted free trees (Cunningham, 1978), and additive similarity trees (Sattath & Tversky, 1977) yield network structures, but the resulting structures must be hierarchical (tree structures). Often this constraint is not appropriate, and the resulting solutions may distort certain relations in the data. The additive clustering method (Shepard & Arabie, 1979) allows for overlapping clusters of entities, which helps avoid the distortions that result from imposing a hierarchical structure on the data. Pathfinder can reveal tree structures in data, but it can also reveal other, more complex, structures that do not obey the hierarchical restriction. Pathfinder can also suggest clusters of entities in the form of interconnected subsets of the entities or cycles in the network (Schvaneveldt *et al.,* 1985).

Another problem of some concern in selecting among scaling methods arises when the proximities are asymmetric, such that the proximity between entities depends on their order. Tversky (1977) has made the case for the psychological reality of asymmetric similarity relations in conceptual organization, and he proposed a set-theoretic feature model that preserves such asymmetries. Similarly, in recognition of the importance of representing asymmetric data, there have been several proposals for scaling asymmetric data in the MDS framework (e.g., Constantine & Gower, 1978; Harshman, Green, Wind, & Lundy, 1982; Krumhansl, 1978). These methods involve separating symmetric and asymmetric components in the data or using spatial density in the resulting MDS configurations in the computation of distance in space.

Given that links in networks can be directed, Pathfinder can naturally represent asymmetric relations between entities. Networks with directed links allow for zero, one, or two links between any two nodes. With two links, the weights may be different. Thus, asymmetry can be represented by having a link in only one direction or by having links in both directions with different weights.

Each of the several methods available for scaling proximity data captures certain aspects of the data, often at the sacrifice of other aspects. Many of these methods may be usefully employed together. For example, MDS and Pathfinder used together can simultaneously reveal an underlying dimensional structure in a set of entities as well as the most salient pairwise relations among them. The appropriate choice for a given set of data will depend on a number of factors such as assumptions about the data, the theoretical motivations behind the work, the kind of information needed, and the interpretability of the resulting solutions. Often meeting these criteria will require more than a single scaling method.

Our work in applying Pathfinder to empirical data has made use of several concepts from graph theory such as minimal spanning trees, cycles, and Hamiltonian cycles. Several other concepts from graph theory could prove useful in characterizing the structure of networks. Some examples of these concepts are (1) *median:* the node with minimum distance from itself to all other nodes in the network; (2) *center:* the node with minimum distance from itself to the most distant node in the network; (3) *basis:* the smallest set of nodes from which every node in the network can be reached; and (4) *minimal dominating node set:* the smallest set of nodes such that every node in the network is connected to a node in the set with one link.

These properties of networks have proven useful in various applications of graph theory, and it would be worthwhile to explore their applicability in the scaling and interpretation of proximity data. We intend to pursue such investigations in further work. Once a network has been determined for a set of data, many quantitative and qualitative properties of the network can be derived. Empirical investigations should help determine which of these properties have value for characterizing the structure of data.

Finally, we should mention some of the work performed by ourselves and others in the general area of knowledge engineering. Pathfinder has proven to be a cornerstone of this work, and future developments of our work in network analysis will be influenced by the needs of these applications.

Roske-Hofstrand and Paap (1986) have used Pathfinder to design a system of menu panels in an information retrieval system used by pilots. The Pathfinder-based system led to superior performance in using the retrieval system by the target users of the system. A similar application of Pathfinder to a menu-based version of the MS-DOS operating system was reported by Snyder *et al.* (1985). Snyder *et al.* reported significantly faster learning of operating system commands with a menu organized according to a Pathfinder network. McDonald and his colleagues (McDonald, Dearholt, Paap, & Schvaneveldt, 1986; McDonald & Schvaneveldt, 1988) have used Pathfinder in conjunction with other scaling methods to design various aspects of the user interface. A major theme in that work is the use of empirical techniques to define users' models of systems. These models are then incorporated into the user interface. Cooke and McDonald (1986) and Schvaneveldt *et al.* (1985) discuss the use of Pathfinder and other scaling techniques in eliciting and representing expert knowledge for use in expert systems. These papers argue that empirically based measurement and scaling procedures have much to offer in the process of defining and codifying the knowledge of experts.

In conclusion, we have been encouraged by the results obtained using Pathfinder networks to identify structure in proximity data. There are also several new avenues to explore in the realm of graph theory that should provide useful structural descriptions. Some of the initial applications of Pathfinder have met with sufficient success to encourage further application and development of the technique. We hope that other researchers will also find Pathfinder a useful addition to the analytic tools available for uncovering latent structure in proximity data.

Acknowledgments

This research was supported by Grant IST-8506706 from the National Science Foundation, by Contract F33615-C-0004 from the Air Force Human Resources Laboratory, by the Computing Research Laboratory, and by an Arts & Sciences summer fellowship from the University of Oklahoma. We thank Tim Breen, Nancy Cooke, and Tim Goldsmith for their contributions to this work. We also appreciate comments on earlier drafts from Michael Friendly, George Furnas, Frank Harary, Jim McDonald, Ken Paap, Keith Phillips, Joe Rodgers, Richard Young, and Mark Wells.

References

Aho, A. V., Hopcroft, J. E., & Ullman, J. D. (1974). *The design and analysis of computer algorithms*. Reading, MA: Addison-Wesley.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Beals, R., Krantz, D. H., & Tversky, A. (1968). Foundations of multidimensional scaling. *Psychological Review,* **75,** 127–142.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology,* **49,** 229–240.

Bower, G. H. (1972). A selective review of organizational factors in memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press.

Butler, K. A., & Corter, J. E. (1986). The use of psychometric tools for knowledge acquisition: A case study. In W. Gale (Ed.), *Artificial intelligence and statistics*. Reading, MA: Addison-Wesley.

Carre, B. A. (1979). *Graphs and networks*. Oxford: Clarendon Press.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology,* **4,** 55–81.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science,* **5,** 121–152.

Christofides, N. (1975). *Graph theory: An algorithmic approach*. New York: Academic Press.

Cohen, B. H., Bousfield, W. A., & Whitmarsh, G. A. (1957). *Cultural norms for items in 43 categories* (Tech. Rep. No. 22; ONR Contract Nonr-631(00)). Storrs: University of Connecticut.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review,* **82,** 407–428.

Constantine, A. G., & Gower, J. C. (1978). Graphical representation of asymmetric matrices. *Applied Statistics,* **27,** 297–304.

Cooke, N. M., Durso, F. T., & Schvaneveldt, R. W. (1986). Recall and measures of memory organization. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **12,** 538–549.

Cooke, N. M., & McDonald, J. E. (1986). A formal methodology for acquiring and representing expert knowledge. *Proceedings of the IEEE,* **110,** 1422–1430.

Coombs, C. H., Dawes, R. M., & Tversky, A. (1970). *Mathematical psychology: An elementary introduction*. Englewood Cliffs, NJ: Prentice-Hall.

Cunningham, J. P. (1978). Free trees and bidirectional trees as representations of psychological distance. *Journal of Mathematical Psychology*, **17**, 165–188.

Dearholt, D. W., Schvaneveldt, R. W., & Durso, F. T. (1985). *Properties of networks derived from proximities* (Memorandum in Computer and Cognitive Science, MCCS-85-14). Las Cruces: New Mexico State University, Computing Research Laboratory.

Droge, U., & Feger, H. (1983). *Ordinal network scaling*. Paper presented at the joint meeting of the Classification Society and the Psychometric Society, Paris.

Ekman, G. (1954). Dimensions of color vision. *Journal of Psychology,* **38,** 467–474.

Feger, H., & Bien, W. (1982). Network unfolding. *Social Networks,* **4,** 257–283.

Fillenbaum, S., & Rapoport, A. (1971). *Structures in the subjective lexicon*. New York: Academic Press.

Friendly, M. (1977). In search of the M-gram: The structure of organization in free recall. *Cognitive Psychology,* **9,** 188–249.

Friendly, M. (1979). Methods for finding graphic representations of associative memory structures. In C. R. Puff (Ed.), *Memory organization and structure*. New York: Academic Press.

Hage, P., & Harary, F. (1983). *Structural models in anthropology*. London: Cambridge University Press.

Harary, F. (1969). *Graph theory*. Reading, MA: Addison-Wesley.

Harary, F., Norman, R. Z., & Cartwright, D. (1965). *Structural models: An introduction to the theory of directed graphs*. New York: Wiley.

Harshman, R. A., Green, P. E., Wind, Y., & Lundy, M. E. (1982). A model for the analysis of asymmetric data in marketing research. *Marketing Science,* **1,** 205–242.

Hutchinson, J. W. (1981). *Network representations of psychological relations*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Isaac, P. D., & Poor, D. D. S. (1974). On the determination of appropriate dimensionality in data with error. *Psychometrika,* **39,** 91–109.

Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika,* **32,** 241–254.

Knoke, D., & Kuklinski, J. H. (1982). *Network analysis*. Beverly Hills, CA: Sage.

Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. *Psychological Review,* **85,** 445–463.

Kruskal, J. B. (1956). On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society,* **7,** 48–50.

Kruskal, J. B. (1964). Nonmetric multidimensional scaling: A numerical method. *Psychometrika,* **29,** 115–129.

Kruskal, J. B. (1977). Multidimensional scaling and other methods for discovering structure. In K. Enslein, A. Ralston, & H. Wilf (Eds.), *Statistical methods for digital computers*. New York: Wiley.

Marshall, G. R., & Cofer, C. N. (1970). Single-word free-association norms for 328 responses from the Connecticut cultural norms for verbal items in categories. In L. Postman & G. Keppel (Eds.), *Norms of word association*. New York: Academic Press.

Maxwell, K. J. (1983). A scripts analysis of fact retrieval from memory. Unpublished doctoral dissertation, New Mexico State University, Las Cruces.

McDonald, J. E., Dearholt, D. W., Paap, K. R., & Schvaneveldt, R. W. (1986). A formal interface design methodology based on user knowledge. *Proceedings of CHI '86,* pp. 285–290.

McDonald, J. E., & Schvaneveldt, R. W. (1988). The application of user knowledge to interface design. In R. Guindon (Ed.), *Cognitive science and its applications for human–computer interaction.* Hillsdale, NJ: Erlbaum.

McKeithen, K. B., Reitman, J. S., Rueter, H. H., & Hirtle, S. C. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology,* **13,** 307–325.

Meyer, D. E., & Schvaneveldt, R. W. (1976). Meaning, memory structure and mental processes. *Science,* **192,** 27–33.

Minsky, M. (1975). A framework for representing knowledge. In P. Winston (Ed.), *The psychology of computer vision.* New York: McGraw-Hill.

Nilsson, N. (1965). *Learning machines: Foundations of trainable pattern-classification systems.* New York: McGraw-Hill.

Quillian, M. R. (1969). The teachable language comprehender: A simulation program and theory of language. *Communications of the ACM,* **12,** 459–476.

Reitman, J. S. (1976). Skilled perception in Go: Deducing memory structures from interresponse times. *Cognitive Psychology,* **8,** 336–356.

Rosch, E., Mervis, C. B., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology,* **8,** 382–439.

Roske-Hofstrand, R. J., & Paap, K. R. (1986). Cognitive networks as a guide to menu organization: An application in the automated cockpit. *Ergonomics,* **29,** 1301–1311.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. Anderson, R. Spiro, & W. Montague (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale, NJ: Erlbaum.

Sattath, S., & Tversky, A. (1977). Additive similarity trees. *Psychometrika,* **42,** 319–345.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding.* Hillsdale, NJ: Erlbaum.

Schvaneveldt, R. W., Dearholt, D. W., & Durso, F. T. (1988). Graph theoretic foundations of Pathfinder networks. *Computers and Mathematics with Applications,* **15,** 337–345.

Schvaneveldt, R. W., & Durso, F. T. (1981). *General semantic networks.* Paper presented at the annual meeting of the Psychonomic Society, Philadelphia.

Schvaneveldt, R., Durso, F., Goldsmith, T., Breen, T., Cooke, N., Tucker, R., & DeMaio, J. (1985). Measuring the structure of expertise. *International Journal of Man-Machine Studies,* **23,** 699–728.

Shepard, R. N. (1962a). Analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika,* **27,** 125–140.

Shepard, R. N. (1962b). Analysis of proximities: Multidimensional scaling with an unknown distance function. II. *Psychometrika,* **27,** 219–246.

Shepard, R. N. (1974). Representation of structure in similarity data: Problems and prospects. *Psychometrika,* **39,** 373–421.

Shepard, R. N., & Arabie, P. (1979). Additive clustering: Representation of similarities as combinations of discrete overlapping properties. *Psychological Review,* **86,** 87–123.

Snyder, K., Paap, K., Lewis, J., Rotella, J., Happ, A., Malcus, L., & Dyck, J. (1985). *Using cognitive networks to create menus* (Tech. Rep. No. TR 54.405). Boca Raton, FL: IBM.

Stevens, S. S. (1951). Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley.

Tversky, A. (1977). Features of similarity. *Psychological Review,* **84,** 327–352.