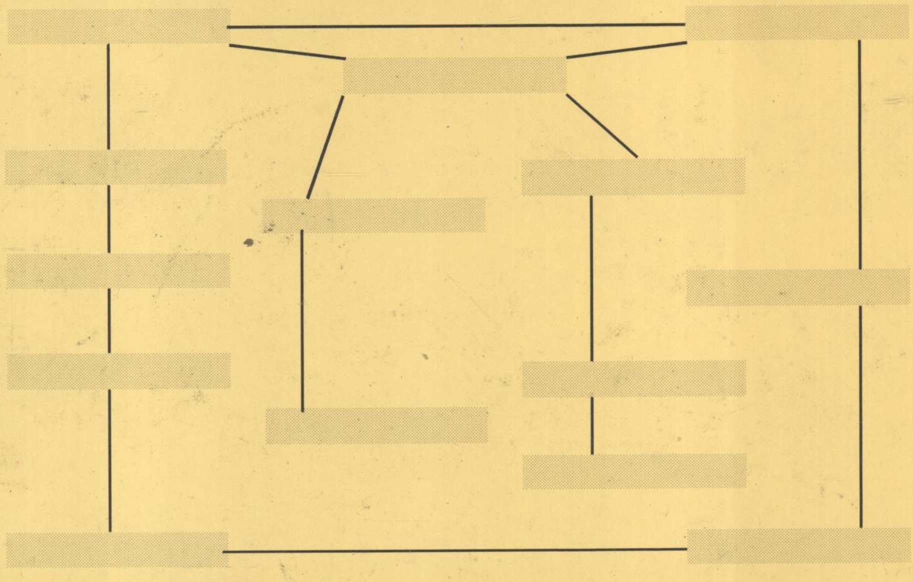
Pathfinder Associative Networks

Studies in Knowledge Organization



edited by

Roger W. Schvaneveldt

Pathfinder Associative Networks: Studies in Knowledge Organization

Edited by

Roger W. Schvaneveldt
Department of Psychology and
Computing Research Laboratory
New Mexico State University

ABLEX PUBLISHING CORPORATION
Norwood, New Jersey 07648

Printed in the United States of America.

Library of Congress Cataloging-in-Publication Data

Pathfinder associative networks : studies in knowledge organization / edited by Roger W. Schvaneveldt.

p. cm. – (Ablex series in computational sciences)
Includes bibliographical references.
ISBN 0-89391-624-2
1. Expert systems (Computer science). 2. Artificial intelligence. 3. Graph theory.
4. Knowledge acquisition (Expert systems).
I. Schvaneveldt, Roger W. II. Series.
QA76.76.E95P28 1989
006.3'3–dc20 89-18219
CIP

ABLEX Publishing Corporation
355 Chestnut Street
Norwood, New Jersey 07648

Pathfinder Associative Networks: Studies in Knowledge Organization

ABLEX SERIES IN COMPUTATIONAL SCIENCES

Derek Partridge, University of Exeter

Series Editor

Pathfinder Associative Networks: Studies in Knowledge Organization
Roger W. Schvaneveldt (Editor)

In Preparation

A New Guide to Artificial Intelligence
Derek Partridge

Artificial Intelligence and Software Engineering
Derek Partridge (Editor)

Binding Time – Six Studies in Programming Technology and Milieu
Mark Halpern

Artificial Intelligence and Business Management
Derek Partridge and K.M. Hussain

Intelligent Systems
Eric Dietrich and Chris Fields

Database in Practice Rather Than in Theory
K.G. Jeffrey

New Generation Architectures and Languages
Stephen J. Turner

Computer Analysis of English: Lexical Semantics and Preference Semantics Analysis
Brian Slator

Contributors

Russell J. Branaghan
Department of Psychology and
Computing Research Laboratory
New Mexico State University
Las Cruces, NM 88003

Timothy J. Breen
Knowledge Systems Laboratory
Boeing Computer Services
P.O. Box 24346, MS 7L-64
Seattle WA 98124-0346

Kathy A. Coggins
Department of Psychology
University of Oklahoma
Norman, OK 73019

Nancy Jaworski Cooke
Department of Psychology
Rice University
Box 1892
Houston, TX 77251

Daniel M. Davenport
P. O. Box 5800
Sandia National Laboratories
Albuquerque, NM 87185

Tom Dayton
Department of Psychology
University of Oklahoma
Norman, OK 73019

Donald W. Dearholt
Department of Computer Science and
Computing Research Laboratory
New Mexico State University
Las Cruces, NM 88003

Francis T. Durso
Department of Psychology
University of Oklahoma
Norman, OK 73019

Chris Esposito
Boeing Advanced Technology Center
P.O. Box 24346 MS 7L-64
Seattle, WA 98124

Richard H. Fowler
Department of Mathematics and
Computer Science
Pan American University
Edinburg, TX 78539

John G. Gammack
MRC Applied Psychology Unit
15 Chaucer Road
Cambridge, CB2 2EF
England

Timothy E. Goldsmith
Department of Psychology
University of New Mexico
Albuquerque, NM 87131

Peder J. Johnson
Department of Psychology
University of New Mexico
Albuquerque, NM 87131

Wendy A. Kellogg
User Interface Institute
IBM Thomas J. Watson Research
Center
P. O. Box 704
Yorktown Heights, NY 10598

Deborah R. McDonald
Department of Psychology
New Mexico State University
Las Cruces, NM 88003

James E. McDonald
Department of Psychology and
Computing Research Laboratory
New Mexico State University
Las Cruces, NM 88003

Lisa A. Onorato
Department of Psychology
Hartwick College
Oneonta, NY 13820

Kenneth R. Paap
Department of Psychology and
Computing Research Laboratory
New Mexico State University
Las Cruces, NM 88003

Tony A. Plate
Department of Computer Science
University of Toronto
Toronto M5S 1A1
Canada

Renate J. Roske-Hofstrand
Aerospace Human Factors Research
Division
Nasa-Ames Research Center
Mailstop 239-21
Moffett Field, CA 94035

David C. Rubin
Department of Psychology
Duke University
Durham, NC 27706

Roger W. Schvaneveldt
Department of Psychology and
Computing Research Laboratory
New Mexico State University
Las Cruces, NM 88003

Jack D. Shepard
Department of Psychology
University of Oklahoma
Norman, OK 73019

Preface

The chapters in this book represent a sampling of theoretical, empirical, and applied work with *Pathfinder networks*. These networks began in 1981 (Schvaneveldt & Durso, 1981) as an attempt to develop a *network* model for *proximity* data. The intervening years have seen several developments of that original work. A theoretical paper relating Pathfinder networks to fundamental concepts in graph theory (Schvaneveldt, Dearholt, & Durso, 1988) grew out of a conference organized by Frank Harary and Keith Phillips. The chapters in this book represent a wide range of applications for network models.

The original motivation for developing Pathfinder grew out of our realization that although network representations abound in theoretical work in cognitive psychology and artificial intelligence, there were few methods for arriving at a network representation from empirical data. Proximity data offer a convenient starting point for networks. Indeed, proximity data serve as the building block for several interesting structural models such as multidimensional scaling (MDS) and cluster analysis. Essentially, Pathfinder networks are determined by identifying the proximities that provide the most efficient connections between the entities by considering the indirect connections provided by *paths* through other entities. The resulting networks have several interesting properties (see Chapter 1), and they have also proven to be useful in a variety of applications. There are now various algorithms for deriving PFNETs in several computer languages running on several different computers.[1]

There are a few features of this book that should be helpful to readers without much background knowledge of graph theory. A brief primer on graph theory and Pathfinder and a glossary can be found at the back of the book. References from all of the chapters are compiled in a single reference section at the back of the book. Chapter 1 reviews some definitions and properties of Pathfinder networks as well as some algorithms for deriving these networks from proximity data. Chapter 2 is a general review of empirical work with Pathfinder in cognitive modeling and an exploration of potential applications in social networks. The other chapters relate to several major themes.

Chapters 3, 4, and 5 address some methodological issues. Esposito (Chapter 3) develops and evaluates a version of Pathfinder that takes variability of proximity data into account. Roske-Hofstrand and Paap (Chapter 4) analyze some properties of proximity data obtained by ratings and the implications for Pathfinder networks. Goldsmith and Davenport (Chapter 5) present some measures of the similarity of two networks.

Chapters 6 through 10 report investigations of some basic phenomena in human memory. Esposito (Chapter 6) analyzes the relation between human judgments of the goodness of categories and various formal characteristics of graphs. Cooke (Chapter 7) examines the time required to judge that two concepts belong to the same category. Branaghan (Chapter 8) analyzes the ease with which lists of associations are learned. Rubin (Chapter 9) investigates the strategies people use to search memory. Schvaneveldt (Chapter 10) examines the representation of schemata in Pathfinder and connectionist style networks.

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

Chapters 11 through 16 address applications of Pathfinder to problems in knowledge elicitation, information retrieval, and interface design. McDonald, Plate, and Schvaneveldt (Chapter 11) extract associative information from text and use this information to resolve word ambiguity. Fowler and Dearholt (Chapter 12) address the classic problem of retrieving information from large collections as in libraries. Kellogg and Breen (Chapter 13) compare the models of systems to mental models of users. McDonald, Paap, and McDonald (Chapter 14) attack the problem of establishing connections in Hypertext. Gammack (Chapter 15) analyzes the use of different techniques for eliciting proximity information from an expert. Cooke (Chapter 16) develops a method for identifying the nature of the relations between linked concepts in a network.

Chapters 17, 18, and 19 are concerned with still other aspects of knowledge representation. Goldsmith and Johnson (Chapter 17) investigate the use of networks and MDS spaces to assess classroom learning. Onorato (Chapter 18) analyzes the ways in which people organize information depending on the purpose of the information. Dayton, Durso, and Shepard (Chapter 19) examine the differences in the way solvers and nonsolvers organize problem-relevant information.

Obviously, there are many interrelations among the various chapters. As an aid to seeing these relations and as an initial illustration of the use of Pathfinder, I constructed Figure 1. This figure shows a Pathfinder network depicting the close associations among the chapters.

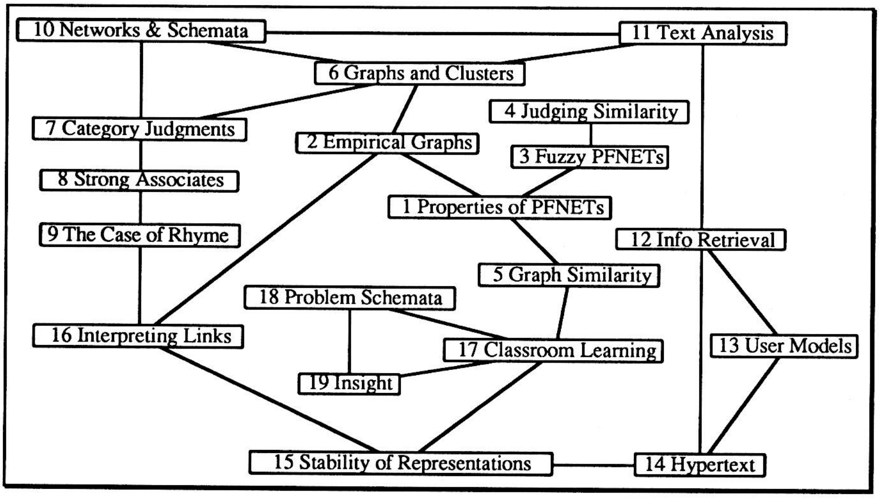


Figure 1. The modified PFNET($r = \infty$, $q = n-1$) for the chapters in this book.

To construct Figure 1, I first made a list of the three chapters most closely related to each chapter. This ordered list of associations was used to construct a matrix of proximity data where an entry was 1, 2, or 3 if the chapter on the column was the first , second, or third most associated with the chapter on the row. Other entries were treated as infinite. This matrix was non-symmetrical and the Pathfinder network that resulted from analyzing the matrix had directed links. However, I was not able to interpret the directions of the links so I made all of the links undirected as shown in the figure.

This figure can be used to find chapters that are closely related to other chapters. Several groups of interrelated chapters can be identified in addition to the one I used to order the chapters. It is obviously impossible to capture all of these relations in the linear ordering enforced on a book.

The development of Pathfinder and much of the research reported in this book have been supported by the National Science Foundation (IST-8506706), the Air Force Human Resources Laboratory (F33615-84-C-0072 and F33615-80-C-0004), Texas Instruments, Inc., and the National Aeronautic and Space Administration (NAG 2-453). Such support has been invaluable in the development of the methods and research.

I gratefully acknowledge the assistance of several people in assembling this book. Derek Partridge encouraged me to undertake the project in the first place. Most of the authors reviewed one or more chapters in addition to writing their own. Douglas Nelson and David Farwell also provided very helpful reviews. My associates here at New Mexico State were invaluable in their assistance with all of the details and in the defense of conceptual coherence. I am particularly grateful to Bob Fiegel, Tarra Fiegel, Rebecca Gomez, and Paula Moreland for their help. Thanks also to my wife, Ann, and daughter, Susan, for their love and support.

R. Schvaneveldt
April 11, 1989
Las Cruces, New Mexico

Contents

Chapter 1

Properties of Pathfinder Networks

Donald W. Dearholt and Roger W. Schvaneveldt

Network models have played important roles in various areas of cognitive science and computer science. In cognitive psychology and artificial intelligence, network representations of concepts stored in semantic memory have been used in models of memory retrieval and human performance (e.g., Anderson, 1983; Collins & Loftus, 1975; Collins & Quillian, 1969; Friendly, 1977, 1979; Meyer & Schvaneveldt, 1976; Rumelhart & McClelland, 1986), scene description and analysis (e.g., Brooks & Binford, 1980; Waltz, 1972; Winston, 1970), natural-language processing (e.g., Bobrow & Webber, 1980; Fillenbaum & Rapoport, 1971; Kintsch, 1974; Quillian, 1967, 1969; Schank, 1972; Woods, Kaplan, & Nash-Webber, 1972), and knowledge representation (e.g., Brachman, 1977, 1979; Fahlman, 1979; Fikes & Hendrix, 1977; Griffith, 1982; Novak, 1977; Schmolze & Lipkis, 1983; Sowa, 1984; Woods, 1975).

In database systems, a network data model often results in efficient representations of sets of concepts (Date, 1981; Ullman, 1982). Thus far, the network model incorporated in database systems has been constructed with two primary objectives: providing efficient data access for the anticipated user environment and making the most of the rather severe limitations imposed by present computer operations and architecture. Although the network data model used most frequently in database models (CODASYL, 1971) can support abstractions of essentially any type, there are constraints (for the purpose of modularity, simplicity of definition, and hardware support) that must be circumvented by artificial programming devices. Networks identifying relationships between data items have been proposed for designing the logical schema of a database system (e.g., see Martin, 1977, Chapter 6) by means of bubble charts. The bubble charts are used to indicate relationships between data items (e.g., functional dependencies, primary keys, and secondary keys). The bubble charts are usually viewed as an intermediate step in the development of a logical schema. Clustering strategies for data items have been investigated and proposed for improving expected retrieval time, based on the estimated likelihood of retrieval of data items contingent upon the retrieval of other data items (e.g., Navathe & Fry, 1976; Schkolnick, 1977).

Recently developed techniques from our laboratory and elsewhere allow researchers to derive networks from the same proximity data employed by multidimensional scaling (Dearholt, Schvaneveldt, & Durso, 1985; Hutchinson, 1989; Schvaneveldt, Dearholt, & Durso, 1988; Schvaneveldt & Durso, 1981; Schvaneveldt, Durso, & Dearholt, 1989).

Networks and Proximity Data

Hutchinson's NETSCAL procedure (Hutchinson, 1981, 1989), which makes *ordinal* data assumptions, is based on a theorem of Hakimi and Yau (1964) regarding the *distance* matrix of a *graph* and its realizability. The distance metric used by Hakimi and Yau is the sum of *edge weights* along a path, so that the distance between *nodes* is the (minimum) sum of the weights (distances) of the edges along a path between the nodes. This measure of *path length* is appropriate for *ratio*-scale data. Hutchinson, however, also used a distance metric in which the distance between nodes is the smallest maximum weight along any of the paths between the two nodes. This path-length measure is appropriate for ordinal as well as ratio-scale data. A serious shortcoming of Hutchinson's work is that his Corollary I considers *triangle inequalities* of only two-link paths. That is, the triangle inequality can be violated in paths having three or more links in Hutchinson's networks. This seems to be an unfortunate limitation, inappropriate for the scaling of data, and perhaps also for cognitive modeling, although psychological proximity may not always obey the triangle inequality (Tversky, 1977; Tversky & Gati, 1982).

The triangle inequality can be viewed in three different domains. The first, and the sparsest, is in euclidean space, as addressed by Hakimi and Yau (1964), in which the triangle inequality must always be satisfied. The second is the class of problems in which measures of similarity or "distance" are measured objectively by set intersections; for most such problems, there is no expectation of transitivity holding, so that there is likewise no expectation that the triangle inequality will be satisfied, either. That is, if we know the intersections of sets *A* and *B*, and of *B* and *C*, we do not generally know anything about the intersection of sets *A* and *C*. The information retrieval application to be discussed in detail in the chapter by Fowler and Dearholt (Chapter 12, this volume) is an example of such a problem in which the triangle inequality may be violated. The third domain is that of subjective estimates of similarity, in which data frequently show violations of the triangle inequality. Philosophically, it is attractive to use geodetic distance measures, in which the distance between each pair of nodes is considered to be the length of the shortest path available between those nodes; indeed, in graph theory this has been the usual definition of distance. Then, a violation of some triangle inequality is never a part of a path between any pair of nodes, because a shorter alternative path is always available. Thus the omission of the edges which violate some triangle inequality in a network assures the preservation of the (geodetic) distances between all pairs of nodes and provides a simpler structure which possesses precisely those edges which are responsible for the most economical paths (Schvaneveldt, Dearholt, & Durso, 1988).

The links that are omitted include those due to differences on two or more separable dimensions, in which the triangle inequality is expected to be violated, as discussed by Tversky and Gati (1982). That is, if *A* and *B* are judged to be similar because of feature *x*, and *B* and *C* are judged to be similar because of feature *y*, then *A* and *C* will normally be judged to be less similar than the triangle inequality would indicate. Thus if salient associations are linked in a graph paradigm, the absence of a *link* can denote a difference in the basis for judged similarity.

We have developed a procedure called Pathfinder (several equivalent procedures, actually) to generate a class of networks called PFNETs, which are based on estimates or measures of distances between pairs of entities. This procedure allows a spectrum of assumptions to be made about the data, including ordinal and ratio properties. The data required are either similarities or distances. Similarities can be obtained either from a subject's estimates of the similarity of each pair of entities in the set or from a measure of set intersections. Distances can be obtained in some domains by estimating or computing appropriate differences between all pairs of entities. The result of the Pathfinder procedure is a network which is either a directed graph (if the similarity or distance matrix is not symmetrical) or an undirected graph otherwise. Each entity in the set is represented as a node in the network, and each link that is entered in the network has a weight value determined by the distance between the two entities so linked. Our network generation procedure incorporates two parameters. The first, the Minkowski r-metric, determines how distance between two nodes not directly linked is computed. The weight of a path with weights w_1, $w_2, \ldots, w_k$ is:

$$W(P)=\left(\sum_{i=1}^{k} w_i^r\right)^{1/r}$$

For $r = 1$, the path weight is the sum of the link weights along the path; for $r = 2$, the path weight is computed as euclidean distance is computed; and for $r = \infty$, the path weight is the same as the maximum weight associated with any link along the path. We will use "distance" in this chapter to mean the Minkowski distance (geodetic), which depends upon the value of the r-metric. The second parameter is the *q parameter*, which is a limit on the number of links in the paths examined in constructing a network. Its value determines the maximum number of links in paths in which the triangle inequalities are guaranteed to be satisfied in the resulting network. Our procedure generates families of PFNETs, and we can generate Hutchinson's (1989) networks as a special case with $r = \infty$ and $q = 2$.

The links omitted from a PFNET are omitted because they violate a triangle inequality involving q or fewer links. These omissions preserve all (geodetic) distances from the original data, however, and because not all links are present in most PFNETs, structural features are easier to ascertain. If a distance between two nodes not directly linked must be computed from the PFNET, it is computed using the Minkowski metric, resulting in a computed distance less than that given explicitly in the original data.

Advantages of PFNETs include (1) the capability of directly modeling asymmetrical relationships (Hutchinson, 1981; Tversky, 1977), which is more difficult with multidimensional scaling (Constantine & Gower, 1978; Harshman, Green, Wind, & Lundy, 1982; Krumhansl, 1978); (2) the provision for a complementary alternative to multidimensional scaling which often provides a more accurate representation of local data relationships than does multidimensional scaling; since multidimensional scaling must move data points to minimize a global error criterion, the resulting relationships between neighboring points is often significantly different than the original data would indicate; (3) the fact that hierarchical constraints in most cluster analysis techniques do not apply to PFNETs; (4) the representation of the most "salient" relationships present in the data; (5) the provision for a new paradigm in studying models of classification; and (6) the provision for a more quantitative paradigm for some of the issues in which networks have been invoked qualitatively or designed intuitively.

From the viewpoint of cognitive modeling, a disadvantage of PFNETs in the present state of development is that we have no way of knowing the features upon which similarity judgments are made. Thus the semantic content of links is not easily discernible (but see Cooke, Chapter 16, this volume). The empirical data we have collected, however, should be viewed as similarity estimates having components which may be unknown; but the use of such data seems important in bridging the gap between the more standard semantic networks (in which the researcher labels links according to his preferences or beliefs at the time) and a more objective representation of the knowledge of interest. For domains in which objective measures of distance are available, PFNETs provide unique representations of underlying structure not obtainable from any other scaling method.

Definitions and Alternatives

In this section we present definitions to provide the proper foundation for the generation procedures and theorems that follow. A PFNET has n nodes, denoted $N_1, N_2, \ldots, N_n$ (or $N_a, N_b, \ldots$). A *link* is an association between a pair of nodes which can be either undirected or directed. A *directed link* is called an *arc*, and an *undirected link* is called an *edge*. In this chapter we will deal mainly with networks having undirected links, or edges, but some of the definitions are more general, and a few examples of directed networks will be given to illustrate this generality. In this spirit, links are labeled e_{ij}, for the edge between N_i and N_j (or for the arc from N_i to N_j). N_i and N_j are *end nodes* of the link e_{ij}. The distance from node N_i to N_j (along the link e_{ij}) is the weight w_{ij}, and these weights are often written in matrix form as an $n \times n$ matrix W. The elements of W can be considered as distances between nodes along the direct paths between every pair of nodes. The distances are often considered as dissimilarities, and W is called either the *adjacency* or *weight* matrix. We assume that $w_{ii} = 0$ and $w_{ij} > 0$ for $1 \le i, j \le n$ where $i \neq j$. If this matrix is symmetric, then a PFNET derived from it is an undirected network. Typically the distance measures (weights) for each pair of entities (nodes) are found either empirically, from similarity estimates by human subjects, or analytically, using some appropriate measure of set intersection and set union, or some distance metric between entities.

A path from node N_a to node N_e, passing through nodes N_b, N_c, and N_d, is denoted by P_{abcde} (if the intermediate nodes are important) or P_{ae} otherwise. The former presumes the existence of edges e_{ab}, e_{bc}, e_{cd}, and e_{de} (either undirected or with appropriate directions), whereas P_{ae} presumes the existence of some unspecified set of edges (or arcs) connecting N_a and N_e. The weight of a path P is denoted $W(P_{ae})$, and the function $W(P)$ is determined by the r-metric and the weights w_{ij}.

The triangle inequality is incorporated into our generation procedure by means of the q parameter.

Definition 1

A network is q-triangular if and only if all possible triangle inequalities involving paths with $m \le q$ links are satisfied, using links and weights in the graph and the r-metric chosen. An example is the triangle inequality

$$w_{ae} \le \left(w_{ab}^{r} + w_{bc}^{r} + \ldots + w_{de}^{r} \right)^{1/r}$$

which is a constraint on the weights of two alternate paths between nodes N_a and N_e. For a graph with n nodes, there can be at most $n-1$ edges in any path in which there is no *cycle*. Thus the q parameter is at most $n-1$. Geodetic distances in the network are unchanged if edges which would violate triangle inequalities are omitted.

Definition 2

The (geodetic) network distance d_{ij} between nodes N_i and N_j is computed as a function of all path weights $W(P_{ij})$, for all paths P_{ij} which connect nodes N_i and N_j as

$$D_{ij} = \mathrm{MIN}\left(W(P_{ij1}), W(P_{ij2}), \ldots, W(P_{ijm}) \right)$$

That is, the distance between two nodes is the weight of the smallest path between those nodes, with all path weights calculated using the (same) appropriate r-metric.

The r-metric and the q parameter provide the elements needed to assure that the networks generated from a particular set of proximity data possess the metric properties discussed in Hakimi and Yau (1964), with the following provisions:

1. The distance from a node to itself is assumed to be zero.
2. The data matrix must be symmetric so that the PFNET is undirected; then the distance between any pair of nodes is independent of direction.
3. The triangle inequality is satisfied for all paths having as many as q edges. To assure that no triangle inequalities whatsoever are violated, q can be set to the number of nodes less one.

For situations in which these metric axioms are satisfied, the concept of distance along a path is the same as the weight of that path. Because the r-metric can take on values from one through infinity, and the q parameter can take on values from one through the number of nodes less one, many different PFNETs can be constructed from a given set of proximity data. However, different values of r and q can result in the generation of the same (*isomorphic*) PFNETs. Frequently, important information from a given set of proximity data can be obtained from different PFNETs, constructed using different values of r and q. Thus it is often not essential that particular choices for r and q be made, to the exclusion of other values. Furthermore, it is sometimes desirable (in cognitive modeling, for example) to violate the metric axioms presented above (also in Hakimi & Yau, 1964; and in Tversky & Gati, 1982). The possibility of constructing directed PFNETs from asymmetric proximity data and (independently) of varying the q parameter provide ways of violating these axioms which correspond to observations about human performance (see, for example, Ortony, 1979; Tversky, 1977; and Tversky & Gati, 1982). Modeling traffic flow on one-way streets provides another example in which asymmetric data are relevant.

Definition 3

A *PFNET(r, q)* is a septuple (*N, E, W, LLR, LMR, r, q*) in which:

N is the set of nodes (concepts), denoted N_i;

E is a square matrix representing names of links in the *complete graph* (i.e., e_{ij} is the name of the link connecting nodes N_i and N_j);

W is the square weight matrix, and its entries are the weights associated with the links in the corresponding positions of the E matrix. The weights on the main diagonal are assumed to be zero, and the remaining weights are assumed to be finite and nonnegative. Thus w_{ij} is the weight of link e_{ij};

LLR, the link-labeling rule, is the procedure used to determine a label for each link, according to some classification scheme;

LMR, the link-membership rule, is the procedure used to determine whether or not each element of the E matrix is added to the PFNET(r, q);

r is the value of the r-metric, and $1 \le r \le \infty$;

q is the value of the q parameter, and $q \in \{1, 2, ..., n-1\}$, where n is the number of nodes.

Definition 4

The link-membership rule (LMR) for PFNETs (either directed or undirected) is given by the following procedure:

1. Define a network consisting of all nodes (concepts) N_i, but no links;
2. Order all elements e_{ij} of the E matrix in some nondecreasing order of their associated weights w_{ij};
3. Consider each e_{ij}, and include e_{ij} in the PFNET(r, q), if and only if e_{ij} provides a path from N_i to N_j which has a weight at least as small as the weight of any other path having no more than q links, using the r-metric to compute the weights of multiple-link paths.

This definition is useful primarily in establishing the concepts associated with Pathfinder networks; computationally efficient algorithms for generating PFNETs will be given in the next section. As an example of the LMR, consider the weight matrix:

$$W = \begin{matrix} 0 & 1 & 4 & 5 \\ 2 & 0 & 2 & 4 \\ 1 & 4 & 0 & 1 \\ 5 & 3 & 1 & 0 \end{matrix}$$

for the nodes N_1, N_2, N_3, and N_4. The complete graph is as shown in Figure 1. The arcs are not labeled because we have not yet developed a labeling rule for directed PFNETs. (Labeling edges with some LLR does not affect the edge membership of an undirected PFNET, because the edges are put there by the LMR, which makes no use of edge labels.)

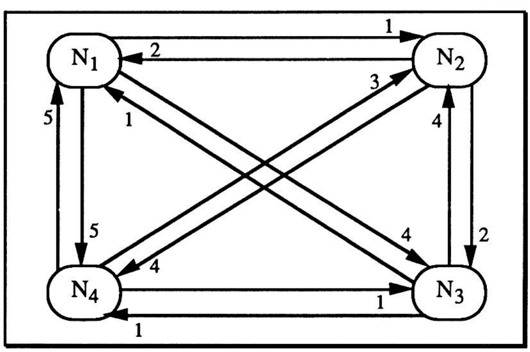


Figure 1. Complete graph for the example.

Let $r = 1$ and $q = 2$. Applying the link membership rule, the PFNET($r = 1$, $q = 2$) shown in Figure 2 is obtained. Note that e_{14} is in the PFNET because its weight ties with the weight of the path P_{124}, even though the arc e_{24} is not itself in the PFNET; if it were in the PFNET, it would violate the triangle inequality for the alternative path P_{234}. The path P_{1234} has less weight, but is not considered because it has three arcs, and for this example we assumed $q = 2$. The PFNET in Figure 2 is two-triangular, since the q parameter is two.