Journal of Classification 11:251-252 (1994)

BOOK REVIEWS

H. Bandemer and W. Näther, *Fuzzy Data Analysis,* Dirdrecht: Kluwer Academic Publishers, 1992, pp. xii+341.

This book is about the use of fuzzy set theory in data analysis. The authors describe fuzziness as being mainly concerned with descriptive data analysis (and hence answering questions of the kind 'What has happened?', where attention is focussed solely on the actual outcome obtained) in contrast to probabilistic theories which, they say, lead to a normative theory, concerned with 'What will happen?'

Chapter 2 reviews basic concepts, defining fuzzy sets, operations on them, and relationships between them.

Chapter 3 describes standard statistical approaches to data analysis. It is, however, a very idiosyncratic overview (e.g., p. 63: 'Transformation of the data is always recommended. The aim of such a transformation...consists in better conditions for recognising not only suspected outliers, but also clusters and functional relationships.'). In discussing distance measures, 40 different ones are listed, but with virtually no critical comment on them. The sections of this chapter cover grouping and transformations, similarity and distances, cluster analysis, evaluation of functional relationships, and projection techniques. Since the penultimate of these is covered in just two and a half pages, only a very cursory and selective outline is given: there is much more to data analysis than is given here!

Chapter 4 describes the sorts of functions that can be used to represent the fuzziness of statements of magnitude and simple operations on fuzzy data.

Chapter 5 deals with what the authors term 'qualitative analysis.' By this they mean (p. 121) that the features of the objects have only two levels

---

Reviewer's Address: D. J. Hand, Statistics Department, The Open University, Walton Hall, Milton Keynes MK7 6AA, England.

(present/absent, true/false, etc.) but allow in this chapter the objects to have degrees of possession of a property. The emphasis here is on fuzzy similarity (including shape similarity) and clustering.

Chapter 6 describes the analysis of fuzzy data in quantitative situations (p. 185): 'this is the same problem as aimed at in statistical inference with crisp data.'

Finally, Chapter 7 is concerned with evaluating methods of fuzzy data analysis. Section 7.1 of this chapter is especially interesting because it discusses the difference between fuzzy theory and probability theory.

The book is aimed at readers who must make inferences or decisions based on experts' opinions or observational data - a fairly broad audience! A working knowledge of mathematics (basic analysis and calculus) and statistics is assumed, but other terms are defined. The authors stress that it is not a 'cookery book,' pointing out that 'since fuzzy set theory derives its application power from its close relation to the given problem and context, the presented and demonstrated methods of data analysis may stimulate the reader to recognize situations and problems in his (sic) environment for their application and to inspire him (sic) how to adapt and generalize these methods.' They also suggest that the book may be useful as a text for graduate students.

Examples can substantially assist in the application of new methods. The authors are aware of this but point out that there are difficulties here since applications of fuzzy set theory depend critically on the context. They have therefore presented simplified examples with references to other work. To me this does not seem to have worked very well, and I would have preferred some more extensive and real examples, even if the penalty was a longer book.

The book certainly gives a good overview of the literature. Unfortunately, the English is not perfect, a fact which can occasionally lead to some ambiguity in meaning and often makes it difficult to follow. With better editing of the grammar and style, with more discussion of the contrasts of statistical approaches, and with some larger and more realistic examples, this could have been a very good book. As it is, it is not bad.

The Open University D. J. Hand

Journal of Classification 11:253-254 (1994)

R. A. Brualdi and H. J. Ryser, *Combinatorial Matrix Theory, Encyclopedia of Mathematics and Its Applications Vol. 39,* Cambridge: Cambridge University Press, 1991, pp. 367.

In the Preface Richard Brualdi writes that in 1984 he proposed to Herb Ryser that they should write together a book on the subject of combinatorial matrix theory. Ryser was greatly intrigued by the idea, and they started the work but never got the chance to finish it together. After Ryser's death in 1985, Brualdi went on with the project, and his interests and views on what is basic material have influenced the selection of topics.

The theme of the book is the symbiotic relationship between matrix theory and combinatorics. The text provides ample illustrations of the beneficial impact of matrix theory and linear algebra on combinatorics as well as of the useful exploration of combinatorial ideas and reasoning in the analysis of matrices. A lot of illustrative examples are chosen so that the reader can feel and share the author's love for the subject. There is in the book a great deal of graph theory, matrix decompositions, network flow theory, and identities involving determinants and permanents of matrices. Many related topics that are omitted or only touched upon in the text are postponed for a planned second volume on Combinatorial Matrix Classes.

As might be evident from this very superficial description of the contents, the book is neither a standard text on matrix theory nor a conventional presentation of graph theory and combinatorics. It is rather a piece of art reflecting the master's way of collecting and organizing some variations of a research theme he has been interested in and contributing to. It reveals mathematical beauty and methodological power. It is likely to attract specialists and to have a great deal of impact on future research.

The first chapter (about 20 pages) on incidence matrices gives fundamental concepts on (0, 1)-matrices, discusses configurations of subsets, isomorphism, the minimax-theorem of König, systems of distinct representatives, permutation matrices, set intersections, and finite projective planes. Chapters 2, 3, and 4 discuss matrix theory for graphs (about 30 pages), digraphs (about 50 pages), and bipartite graphs (about 40 pages). The adjacency matrix, the incidence matrix, the oriented incidence matrix, and the Laplacian matrix are investigated in some detail. Their power as analytical

---

Reviewer's Address: Ove Frank, Department of Statistics, Stockholm University, S-106 91 Stockholm, Sweden.

tools is illustrated by various results on spectral properties, unimodularity, connectivity, complexity, chromatic properties, matchings and factorings. Chapter 5 on some special graphs (about 20 pages) is concerned with regular and strongly regular graphs. Existence theorems is the topic of Chapter 6 (about 30 pages). Here network flows are investigated and several results are given on the existence of matrices with prescribed row and column sums. Chapter 7 (about 50 pages) is devoted to the permanent and its combinatorial usefulness. Chapter 8 (about 40 pages) gives results on latin squares. The final Chapter 9 (about 50 pages) on combinatorial matrix algebra investigates combinatorial properties of determinants and uses algebraic extensions of adjacency matrices and incidence matrices in order to obtain further results on the interplay between matrix algebra and combinatorics.

It is a pleasure to read this book. The style of presentation and the choice of notation is mostly simple and elegant. There are a number of exercises and references following each section. All section references are also given in a general reference list at the end of the book. Sixteen of these references are from 1990 or later and bring you to the frontier of research in this field.

Stockholm University Ove Frank

Journal of Classification 11:255-256 (1994)

Nancy J. Williamson and Michèle Hudon, Eds., *Classification Research for Knowledge Representation and Organization. Proceedings of the 5th International Study Conference on Classification Research, Toronto, Canada, June 24-28, 1991,* Amsterdam: Elsevier, 1992, pp. 427.

The proceedings under review originate from a conference held in Toronto in 1991 which was the 5th in a series all devoted to questions of classification research for application in documentation and indexing systems. The first and second ones, held 1957 in Dorking (1957) and 1965 in Elsinore (Atherton 1965) must now be seen as milestones which presented state-of-the-art surveys and valuable hints for further research and development; they had a vary large influence on the information profession. The next two conferences, 1975 in Bombay (Neelameghan 1979) and 1982 in Augsburg (Dahlberg and Perreault 1982-83), were dominated by the presentation of more specialized papers, a lot of them as a result of a computer application for research in classification and indexing. With this 5th conference proceedings only the process of diversification is perpetuated; except for the title no general theme is to be seen.

For many years classification theory in the library and information science fields was influenced by the construction of classification schemes for representing knowledge systematically in philosophical theories or encyclopedias. Methods for structuring these schemes were dominated by conventional layout in printed schedules. Afterwards the concept of faceted classification schemes was invented to represent complex statements by means of elemental classes and some form of syntax (cf. Buchanan 1979). Nowadays much concern is given to the application of classification schemes in online bibliographical retrieval systems and necessary features for an adequate structure. One is looking for new promises to represent knowledge systematically (whatever this will mean) and classification theorists are interested in similar developments in artificial intelligence or the cognitive sciences. Not all of these inventions are represented in this proceeding volume.

The book is introduced by a keynote address by Elaine Svenonius: "Classification: Prospects, Problems, and Possibilities": it gives a readable

---

Reviewer's Address: W. Gödert, Fachhochschule für Bibliotheks- und Dokumentationswesen in Köln, Claudiusstrasse 1, 5000 Köln 1, Germany.

account of historical traces and philosophical positions but not for further development. The reviewer thinks that this paper is superseded by that of Francis L. Miksa. This paper possibly is the most valuable for anyone - especially outside the field of library and information science - who is interested in the underlying destinations for the construction and application of classification systems historically as well as in the contemporary trend. Miksa's paper originates from one of 9 plenary sessions which are supplemented by 30 contributed papers. Almost all, plenary as well as contributed papers, treat special questions separately, no attempt has been made to group together thematically similar papers.

It is not suitable to go into more details. One group of papers consists of reports on developments in established schemes or projects. A second group discusses applications of known techniques to new special fields (some of them devoted to questions of thesaurus construction and application) and a third one is devoted to exploration of new approaches and techniques in the context of knowledge representation (e.g., expert systems).

In reading these conference proceedings one certainly can get the feeling of disappointment if one expects any red ribbon or unified view as had been present in the first conferences on classification research. The conference summary and conclusions consist of one page and were not debated at the conference - once more in contrast to the former ones.

In our days a unified view of classification research no longer seems to exist - the bulk of specialized work only in the library and information science fields is well represented in this proceedings volume. Perhaps we have to look for new unifying principles within, for example, the cognitive sciences?

Polytechnique of Library and Information Science, Cologne

W. Gödert

References

ATHERTON, P., Ed. (1965), *Classification Research: Proceedings of the 2nd International Study Conference on Classification for Information Retrieval, held at Elsimore, Denmark, 14-18.9, 1964,* Copenhagen: Munksgaard.

BUCHANAN, B. (1979), *Theory of Library Classification,* London: Bingley.

DAHLBERG, I., and PERREAULT, J. M., Eds. (1982-83), *Universal Classification I: Subject Analysis and Ordering Systems, Proceedings of the 4th International Study Conference on Classification Research, Augsburg, 18.6-2.7, 1982, Volumes 1-2,* Frankfurt: Indeks.

INTERNATIONAL STUDY CONFERENCE ON CLASSIFICATION FOR INFORMATION RETRIEVAL (1957), Dorking, England, 13-17.5.1957, Proceedings, London: Aslib.

NEELAMEGHAN, P., Ed. (1979), *Ordering Systems for Global Information Networks: Proceedings of the 3rd International Conference on Classification Research, Bombay, 1975,* Bangalore.

Journal of Classification 11:257-258 (1994)

M. Schader, Ed., *Analyzing and Modeling Data and Knowledge*, Berlin: Springer-Verlag, 1992, pp. 346.

This volume contains the proceedings of the 15th annual conference of the "Gesellschaft für Klassification e. V." held in Salzburg in 1991. The book consists of 35 papers, organized by the editor into three sections. All the papers are written in English.

The first section on Data Analysis, Classification, put together 15 papers presenting rather different methods and/or techniques with varying applications. The second section on Data Modeling, Knowledge Processing, includes 9 papers on modeling, representation, storing and processing data knowledge. The third section presents 11 papers describing several Applications, Special Subjects.

Readers of this journal will almost surely be most interested in the first section and part of the third one. Applications are included in almost all the papers and fall into fields as different as archaeology, astrophysics, biology, chemistry, economy, linguistics, marketing, medicine, and social sciences.

Among the papers I especially appreciated "A Clustering Technique for Maximizing ϕ-divergence, Noncentrality and Discriminating Power" by H. H. Bock, "Multidimensional Scaling as a Framework for Correspondence Analysis and Its Extensions" by R. Meyer, "Conception of a Package for Typology Making and Analyzing" by B. G. Mirkin and M. V. Yeriomin, "Classification and Clustering in Spatial and Image Data" by B. D. Ripley, "The MVL (Missing Values Linkage) Approach for Hierarchical Classification When Data Are Incomplete" by M. Schader and W. Gaul, in the first section. I also liked "Challenges for Information Systems: Representation, Modeling and Metaknowledge" by R. Michalski and F. J. Radermacher, "Stylistic Analyses in Archaeology by Means of Correspondence Analysis" by K. Hoilund Nielsen and "The Multilayer Perceptron for Discriminant Analysis: Two Examples" by F. Murtagh. Of course some of these papers display my personal main areas of interest but on the whole they impressed me as being clear and/or concise and/or comprehensive and/or well structured.

Reviewer's Address: H. Bacelar-Nicolau, Laboratório de Estatística e Análise de Dados, Faculdade de Psicologia e Ciências da Educação, Universidade de Lisboa, Alameda da Universidade, 1600 Lisboa, Portugal.

Classifying the papers into just the three parts described above seems slightly forced to me. Papers in the first section could well be split into two classes. On the other hand, there is some natural overlapping among sections which it appears could easily have been reduced by a different way of clustering the papers. Maybe the distinct partition model I have in mind results from the fact that I couldn't use the same aggregation criterion as the editor, "supported by the keywords assigned to each paper," as is pointed out in the preface. In spite of that information none of the papers presents a keyword list.

On the whole I believe this volume provides an interesting and stimulating sample of the field of analyzing and modeling data and/or knowledge, for researchers as well as for users. As mentioned above there are not keywords in each paper, but the book incorporates a table of contents and a subject index.

University of Lisbon H. Bacelar-Nicolau

Journal of Classification 11:259-260 (1994)

P. S. Moharir, *Pattern Recognition Transforms,* Chichester: Research Studies Press (Wiley), 1992, pp. 272.

The book deals with a special class of feature extraction algorithms having certain useful properties. The author analyzes a number of noninvertible nonlinear transforms which allow the pattern to be recognized independently of its position. In these transforms the transform coefficients remain unaltered if the inputs are cyclically, dyadically or even "randomly" permuted or translated.

After a nice short survey of principal problems in statistical pattern recognition (cluster analysis, classification, estimation of probability density functions, feature selection and feature extraction), in the second chapter the author discusses a notion of translation invariance in detail. The author mainly deals with translation of discrete sequences presented usually in the form of numbers on a ratio scale. He defines cyclic and dyadic translations and their generalization - a mixed-radix half-addition transformation based on a mixed radix number system. Further, he analyzes useful properties of this class of transformations and finds circumstances when these transforms are invariant to different types of distortions of the input signal. Pattern recognition is only one type of problem for which such types of transform can be useful.

In the third chapter, the author analyzes the problem of the efficient computations required to perform transformations analyzed in the previous chapter. The advantages of efficiency may not be merely the saving of computational time but rather the possibility of tackling bigger problems. If the transformation kernel has a great deal of structure, symmetry or redundancy, then blocks for a complex transformation can be built from smaller-order ones. A number of structures of the matrices and their products are analyzed and used in order to propose fast computational transformation algorithms.

In the fourth chapter, the concept of a symmetric-function transform is introduced. It allows one to develop and analyze essentially nonlinear transforms and increases the number of transforms having efficient algorithms for translation-invariant pattern recognition. In the fifth chapter, interrelations between different types of transform are studied in order to provide a user with some help in future possible applications.

Reviewer's Address: Sarunas Raudys, Institute of Mathematics and Informatics, Akademijos str 4, 32600 Vilnius, Lithuania.

The book contributes a great deal of new information and a significant generalization concerning a special feature extraction problem. It can be recommended for research workers dealing with signal processing in pattern recognition. However this monograph is not a recipe book: in order to understand different transformations and to choose a proper one for one's specific purposes, one needs to spend some time in order to get used to a great number of abbreviations and notations, concepts of permutations, transformations, etc.

Vilnius, Lithuania Sarunas Raudys

Journal of Classification 11:261-262 (1994)

D. W. Scott, *Multivariate Density Estimation: Theory, Practice and Visualisation,* New York: Wiley, 1992, pp. 317 +8 pages of color illustration.

Over the last few years, there have been several books dealing with topics in density estimation and nonparametric smoothing. Any new text should therefore aim to give a new perspective or cover new areas. This book does so in two ways. First of all, the emphasis is on multivariate problems. This is an ambitious task in view of the well known difficulties of using nonparametric techniques with high dimensional data. However, the author takes the sensible and convincing view that many problems are multivariate in a relatively small number of dimensions, or that where larger numbers of dimensions exist, these can be reduced effectively to a much smaller number. An additional novel slant of the book is its emphasis on visualisation. Simple one and two-dimensional plots are the bread and butter of statistics but here the author uses pseudo three-dimensional plots in an interesting, entertaining and often informative fashion.

The subtitle of the book is "theory, practice and visualisation." In my view, it is the principal strength of the book that these three rather different areas are represented in a balanced way. In the introduction there is a brief reminder of some methods of graphical display for multivariate data but the importance of appropriate functions, rather than observations, is then argued in a convincing manner. It is these functions to which histograms and scatterplots point. This leads into the most interesting and innovative part of the book, which is visualisation. Two-dimensional density estimates can be illustrated effectively and easily through contours. However, contours can also be drawn in three-dimensions. A contour here is a closed surface. These are illustrated in a variety of ways, sometimes with several contours corresponding to different levels superimposed. It can be difficult to communicate shape and structure clearly here, even with sophisticated computer graphics, but the attempt can often be worthwhile.

The basic theory for histograms, frequency polygons and kernel methods is developed next. Practical issues are not forgotten, with discussion of methods of bandwidth selection, computational techniques for large samples and numerous illustrations.

Reviewer's Address: A. W. Bowman, Department of Statistics, University of Glasgow, University Gardens, Glasgow G12 8QW, Scotland.

In the final three chapters, the range of topics is extended considerably into dimensional reduction, nonparametric regression and additive models. Brief pointers are also given to other issues such as classification and bootstrapping. The breadth of coverage of the book is certainly one of its attractions. Some topics, such as histograms, are covered in surprising depth, in that case to lay foundations for the computational benefits of "average shifted histograms." In general, there is plenty of theoretical development for those who wish to pursue this, while the wide variety of examples illustrate the applications of the techniques, particularly visualisation. Exercises are provided at the end of every chapter which would make the book suitable for a graduate text in the area.

In summary, I enjoyed reading this book. It is a well-rounded discussion of a wide area and for anyone who has an interest in this area it will be a very useful book to have around.

University of Glasgow A. W. Bowman

Journal of Classification 11:263-266 (1994)

Brian S. Everitt, ***Cluster Analysis*** (3rd Edition), London: Edward Arnold, 1993, pp. 170.

The first edition of Brian Everitt's book was published in 1974. As a graduate student, I learned much from my study of the first edition. By the time that I had established my own research stream in classification, I was recommending Everitt's book to those who sought out help in applied clustering. My opinion of the most recent edition is positive, but with some reservations. After 15 years of work in this area, I can still recommend the text to those who need to develop a background in classification.

The third edition consists of eight chapters plus an appendix. The first chapter provides an introduction to the field of classification and a helpful discussion of the concept of a cluster. Chapter 2 offers a survey of methods that provide for an initial examination of the data. The chapter presents an overview of a wide range of graphical methods for data evaluation prior to clustering. Chapter 3 reviews proximity measures that are frequently used in classification analyses. Chapters 4 through 7 present the clustering methods themselves grouped into hierarchical, optimization, mixture models, and other techniques. At the end of each chapter, several example analyses are presented from a variety of disciplines. The closing chapter offers some final observations and briefly discusses the problem of comparing classifications.

The appendix offers a review of seven software packages for classification and provides contact information for acquiring the code. To those new to classification, this appendix can be very helpful. Also I was pleased to see an extended index of six pages. I have seen far too many research oriented texts offer a one page index or none at all.

I do have a number of reservations with the third edition of which the reader should be aware. Although the References section contains over 200 entries, I could not find a single citation to an article published in the *Journal of Classification!* After a decade of active publication in the area of classification, it is hard to believe that at least one article of merit has not been published in the *Journal.*

Those citations that are included in the References section do provide reasonable access to the rest of the classification literature. Most major books on classification are cited to aid the reader's search for more information. the

---

Reviewer's Address: Glenn W. Milligan, Department of Management Sciences, 302 Hagerty Hall, The Ohio State University, Columbus, Ohio 43210, USA.

research articles span nearly 30 years of activity in the area. Furthermore, a number of articles are included to illustrate applications of the methodology.

It has been my experience that individuals without a prior course in multivariate methods find previous editions of the book difficult to understand. The Preface to the third edition states that the mathematical level of the text has been kept deliberately low. However, Everitt does not fail to use statistical equations whenever they are needed and introduces matrix notation with no review of the topic. One should not recommend this text to applied researchers without the appropriate statistical background.

As indicated in this listing of chapter topics, the book is heavily oriented towards clustering methods. I believe that the complexity of the overall clustering process has not been fully developed, at least to the point of making the process explicit. In my own literature reviews, I have attempted to develop the view that the clustering method is just one step in an extended sequence of activities that yield a cluster analysis (see Milligan and Cooper 1987, Milligan 1994).

In some situations, I found myself in disagreement with the discussion or recommendations. These areas of disagreement result from my own experience and research into clustering methodology. For example, in Chapter 2, the presentation does a good job at critiquing the deficiencies of several of the graphical analysis methods. However, after presenting Chernoff's Faces, the author failed to note the numerous deficiencies with this approach. In my view, the Faces methodology is far too weak and potentially misleading due to perceptual problems involved in weighting the salient features of each face. I would replace the Faces methodology with a discussion of block-diagonal clustering or other more suitable graphical approach.

One strong feature of Chapter 2 relates to the discussion about the routine use of principal components prior to clustering. Any individual considering this course of action should consider the implications of Figure 2.11 on page 27. A components analysis is likely to obscure or partially hide the groups when the clustering is present in the pre-transformed variable space. This problem is routinely overlooked by those who advocate ordination methods.

A serious weakness in Chapter 3 involves the discussion of variable standardization. Everitt fails to note the results of one of the few evaluative studies on this issue (Milligan and Cooper 1988). I don't take this personally and the omission probably relates more to Everitt's aversion to citing the *Journal of Classification*. However, there appears to be merit to the results of the simulation study. Quite a few individuals conducting applied analyses have contacted me to indicate that recommendations from the study were consistent with their empirical findings.

I also disagree with the position taken by Everitt on page 71 in Chapter 4 regarding the single linkage hierarchical clustering method. Everitt states that "Because of its mathematical advantages and because it can be used on large data sets, it remains a useful technique in some situations." The single link method has performed so poorly in so many simulation studies that it is clear that it has no place in applied classification analysis. The method is so nonrobust to data error and intermediates between clusters that it is of no real benefit in empirical research.

Likewise, it is not clear to me that mixture model clustering methods deserve their own chapter. For more than 15 years, I have been attending various classification conferences. At each conference, presentations have been made that attempt to solve the clustering problem with the mixture model approach. None of these developments have performed with any degree of merit. Although it is tempting to assume multivariate normal mixtures and then try to derive the optimal clustering method, problems such as parameter estimation, among others, quickly overwhelm the analysis. Perhaps this approach to classification will yield significant advances in the future, but it has not done so to date.

I was disappointed to find a rather limited discussion of overlapping clustering methods in Chapter 7. Those methods that are discussed date from 1970 or earlier. From personal experience, I can assure the reader that software is not available from the authors cited in the section on clumping methods. This section would be greatly improved by adding a discussion of the ADCLUS model of Shepard and Arabie (1979) and the OVERCLUS procedure used in SAS (1985).

Finally, the material in Chapter 8 on assessing clustering solutions is not particularly up-to-date, and relies heavily on results obtained prior to 1980 (see Milligan 1994 for a review of more recent work in this area). Similarly, the Fowlkes and Mallows (1983) index is recommended for the comparison of partitions. Although not discussed by the authors, this statistic can result in division by zero for some cluster configurations! Again, better indices are available in the form of the Hubert and Arabie (1985) corrected Rand index. A direct comparison between these measures was conducted by Milligan and Cooper (1986).

Despite these reservations, I do recommend the text by Everitt to those with a background sufficient to read and learn from it. In the future, I plan on handing a copy of this review to each person to whom I recommend the text. In this manner, the reader will benefit from two perspectives of the field of classification.

The Ohio State University Glenn W. Milligan

References

FOWLKES, E. B., and MALLOWS, C. L. (1983), "A Method for Comparing Two Hierarchical Clusterings," *Journal of the American Statistical Association, 78*, 553-584

HUBERT, L. J., and ARABIE, P. (1985), "Comparing Partitions," *Journal of Classification, 2*, 193-218.

MILLIGAN, G. W. (1994), "Clustering Validation: Results and Implications for Applied Analyses," in *Clustering and Classification*, Eds., P. Arabie, L. Hubert, and G. De Soete, River Edge, New Jersey: World Scientific Press, in press.

MILLIGAN, G. W., and COOPER, M. C. (1986), "A Study of the Comparability of External Criteria for Hierarchical Cluster Analysis," *Multivariate Behavioral Research, 21*, 441-458.

MILLIGAN, G. W., and COOPER, M. C. (1987), "Methodology Review: Clustering Methods," *Applied Psychological Measurement, 11*, 329-354.

MILLIGAN, G. W., and COOPER, M. C. (1988), "A STUDY of Variable Standardization," *Journal of Classification, 5*, 181-204.

SAS User's Guide: Statistics, (1985), Cary, NC: SAS Institute.

SHEPARD, R. N., and ARABIE, P. (1979), "Additive Clustering: Representation of Similarities as Combinations of Discrete Overlapping Properties," *Psychological Bulletin, 86*, 87-123.

Journal of Classification 11:267-270 (1994)

T. Terano, M. Sugeno, M. Mukaidono, and K. Shigemasu, Eds., *Fuzzy Engineering toward Human Friendly Systems,* Amsterdam, Washington, and Tokyo: IOS Press/Ohmsha, Ltd, 1992, pp. xxi + 1142.

Fuzzy set theory is receiving more attention, both among technical and nontechnical audiences. The year 1993 saw the introduction of a new scientific journal devoted to fuzzy set research (*IEEE Transactions on Fuzzy Systems*) and a book on fuzzy logic intended for general audiences (McNeill and Freiberger 1993). The increase in attention has been powered in part by the numerous consumer applications developed by the Japanese, such as better performing auto-focus cameras, passenger trains, anti-lock brakes, washing machines, etc. As part of the Japanese research and development effort, the Laboratory for International Fuzzy Engineering Research (LIFE) was created in 1989 to further develop Zadeh's (1965) theory of fuzzy sets and to apply elements of this theory to problems involving computer-based technologies. LIFE and two other groups sponsored the International Fuzzy Engineering Symposium '91 (IFES '91), which was held in Yokohama on November 13-15, 1991. Professor T. Terano, Head of LIFE, identified the objective of the conference as "to disseminate LIFE's results to date, and also to contribute to the development of this field by collecting the latest information in Japan and abroad." *Fuzzy Engineering toward Human Friendly Systems* is the conference proceedings of IFES '91.

The book is organized as:

Introduction (1 paper, 6 pages, by Professor T. Terano)
Special Lecture (1 abstract, 2 pages, by Professor L. Zadeh)
Part I Mathematics (12 papers, 114 pages)
Part II Fuzzy Technology (15 Papers, 160 pages)
Part III Information Processing (7 papers, 72 pages)
Part IV Fuzzy Computing System (4 papers, 48 pages)
Part V Expert System (11 papers, 106 pages)
Part VI Fuzzy and Neuro (9 papers, 96 pages)
Part VII Fuzzy Logic Control (23 papers, 242 pages)

Reviewer's Address: R. J. Hathaway, Mathematics and Computer Science Department, Georgia Southern University, Statesboro, GA 30460-8093, USA.

Part VIII Application (26 papers, 249 pages)
Demonstrations (18 papers, 36 pages)

The introductory section is absolutely fascinating. Fuzzy set theory was introduced to give a rigorous mathematical framework for handling the type of uncertainty that we commonly call vagueness. Humans use vagueness to great advantage, and it is a fact that vagueness is necessary when high levels of complexity are present. The communication problem between man and machine arises because vagueness is not natural to many of the existing information systems with which humans interface. According to Professor Terano, fuzzy engineering can provide us with an interpreter, so that in the future there can be true integration of man and machine to form a "man-machine system." Ideally, this would allow man to occupy himself primarily with right brain thought, while other more left-brained activities are handled by the machine portion of the system. One cannot help but feel some excitement and enthusiasm for the subject after reading Professor Terano's vision of the future.

The two major application areas which have most benefited from the application of fuzzy set techniques are control and pattern recognition, and both are represented to some extent in this book. For those unfamiliar with fuzzy controllers, a very simple example would be a device for monitoring and adjusting an air-conditioning system that implements a rule such as "if the temperature is warm, then slightly increase the cooling lower," where the vague quantities (warm, slight increase) are treated using fuzzy set techniques. The special lecture of Professor Zadeh raises questions about exactly how fuzzy if-then rules can be learned, combined, interpolated, compressed, etc. Control is very well represented in this collection of papers as it is the sole topic of Part VII and the main topic of the last section. Specific control applications include automotive engineering, electric power systems control, hydraulics, helicopter flight, elevator control, and home appliances.

Classification is the main component of only a few of the papers. In one, a fuzzy classification approach is applied to the analysis of time series data corresponding to stock prices. In another, a fuzzy clustering technique is used to restructure a knowledge base so that the corresponding expert system will have increased inference efficiency. There is also a description of a fuzzy-neural scheme for classifying patterns containing a high degree of noise. Another paper introduces a simple neural-based classifier used in an expert system for diagnosing silicosis. A final application uses fuzzy clustering methodology to process a nonstationary signal contaminated by an additive noise.

An assortment of other theoretical and applied topics is mentioned here. Regarding theory, an article by Professor G. Klir gives an interesting

global view of different theories of uncertainty (including probability and fuzzy set theory) using the principle of uncertainty invariance, which is based on the provocative position that "every real-world decision or problem situation involving uncertainty can be formalized in all the theories of uncertainty." Other theoretical papers cover elements of logic, topology and algebra. Applied papers describe various fuzzy-based techniques for seismic modeling, sociogram analysis, exchange trading, telephone network management, financial analysis, and human behavior analysis. Several papers give general techniques for solving fuzzy versions of standard problems such as optimization, matrix inversion, modeling, and prediction. Statistical applications include the use of fuzzy sets as critical regions in the context of hypothesis testing, and a fuzzy procedure to evaluate subjective probabilities. Though not represented accordingly in this review, a large portion of the book is devoted to computer and information science topics such as software and hardware for fuzzy computing, information systems, artificial intelligence, document retrieval, image processing, etc.

Regarding software, one of the papers contains a discussion of the package FUZZYSTAT, which is described as "fuzzy set software for behavioral and social sciences" and is based on the book by Smithson (1987). The capabilities of this software include measurements of various types of fuzzy set overlaps, inclusions, similarities, and degrees of fuzziness. The package also does transformations of fuzzy sets, including negation, concentration, and dilation. The version described is written in FORTRAN and capable of running under UNIX and VMS, and on 68030 based Macintoshes. The address of the author is Dr. Michael Smithson, School of Behavioral Sciences, James Cook University, Queensland 4811, Australia. (The reviewer has not personally used this software.)

Fuzzy Engineering toward Human Friendly Systems is a 1991 sampling of the worldwide efforts to develop and apply fuzzy set theory to a wide variety of problems in industry, medicine, artificial intelligence, socioeconomics, ecology, behavioral science, and education. The international list of authors is outstanding, and the overall quality of most of the papers is good. This collection of papers is recommended to anyone who wants a recent and wide-ranging survey of applied fuzzy set research, particularly in areas such as control and artificial intelligence. This book is not recommended to those interested primarily in fuzzy methodology for clustering and classification.

Georgia Southern University R. J. Hathaway

References

MCNEILL, D., and FREIBERGER, P. (1993), *Fuzzy Logic*, New York: Simon and Schuster.

SMITHSON, M. (1987), *Fuzzy Set Analysis for Behavioral and Social Sciences*, New York: Springer-Verlag.

ZADEH, L. A. (1965), "Fuzzy Sets," *Information and Control, 8*, 338-352.

Journal of Classification 11:271-273 (1994)

R. J. Pankhurst, *Practical Taxonomic Computing*, Cambridge: Canbridge University Press, 1991, pp. 202.

To many biologists, taxonomy, and especially taxonomic computing, is a topic approached with considerable caution; when they are offered a book which suggests a practical approach to dealing with its problems, some may be persuaded that they might begin to delve into its mysteries- they will be disappointed, if not frustrated, by what they find.

This book is an updated version of the author's earlier work (Pankhurst 1978), and is designed to take account of changes which have taken place in the intervening period. These have been not so much in taxonomy as in computer technology and are reflected in the fact that many biologists now have ready access to micro-computers with a power and storage capacity greater than that of many main-frames of the mid-1970's. It has also resulted in a considerable change of emphasis away from the necessity to having to write one's own programs, to a situation where it is feasible to run a range of software packages; a considerable number of these have been developed which are useful to taxonomists.

The extent to which this book discusses this new situation is very variable. A new and useful addition is Chapter 2, on databases. This advocates, very properly, the use of standardized databases for the collection of taxonomic data, and gives a number of examples of how both herbarium collections, and large flora projects, can utilize such techniques. Their relevance to smaller projects is less clear, as there does not appear to be any suitable repository for the data once a project has been completed. The relation of this chapter to the rest of the book is also rather obscure, since it soon becomes apparent that none of the classification or identification packages mentioned subsequently can actually utilize database files without pre-processing. It is also a not particularly well-written chapter, with a significant number of spelling and grammatical errors, and especially the persistent use of 'data' as a singular noun. For some reason, the text is illustrated with a number of terminal commands or program snippets which are irritating to the reader, superfluous to those familiar with the particular software and meaningless to those who are not.

Reviewer's Address: R. B. Ivimey-Cook, Department of Biological Sciences, University of Exeter, Hatherly Laboratories, Prince of Wales Road, Exeter EX4 4PS, England.

A chapter on phenetic classification discusses the information content of characters, and considers clustering, ordination and divisive classification methods. Much of this has been extant for sometime, but it is useful to have it reiterated here. There is also a discussion of phylogenetic and cladistic classification methods, a more controversial topic, though one hardly likely to appeal to those who were not fairly dedicated to taxonomy, and therefore, presumably moderately familiar with the literature. Several software packages for carrying out various facets of classification are mentioned, though in the light of the author's comment that this field is changing rapidly, one must wonder at the relevance of some which are nearly 10 years old.

Identification methods are then discussed. Whether it is really necessary to detail procedures for manual key construction is debatable; I find it difficult to believe that many people will still wish to construct keys from edge-punched cards, and in any event, this seems scarcely relevant to computing. Methods specifically adapted to computing are then discussed, including programs for generating keys from raw data, with which, of course, the author is very familiar. There is a useful account of the use of DELTA for taxonomic descriptions, though this does not seem to be related in any way to the previous chapter on databases.

The book concludes with a brief history of identification methods, some notes on applications and a short chapter on expert systems.

In many ways, this is a most unsatisfactory book. It never seems to have worked out who its readers might be or what is its objective. There is great uneveness in the text. Thus on p. 23: "Consider the name *Bellis perennis* (actually the Common Daisy). This species belongs to the family Compositae, also known as the Asteraceae, but this fact is not explicitly stated in the name." To include the vernacular name is superfluous (and not very meaningful to a non-British reader); as to the second sentence, one presumes that the fact referred to is that this species belongs to the Compositae, though as written, it is that the Compositae is also known as the Asteraceae! In contrast to such rather trite statements, one finds, e.g., on p. 84:

> "The algorithm searches for the largest clique or cliques, since there may be several of the maximum size. The largest clique is then a reasonable choice of characters and this can be used to draw a cladogram without homoplasy, which is the best hypothesis."

which, even within its context, is on a rather different intellectual plane.

Despite its title, I feel it would be extremely difficult for anyone not reasonably familiar with the subject to start a taxonomic computing project with this book as their sole guide; if they were sufficiently familiar with the topics, they would probably not need the book. It may never have been the author's intention that it should be used in such a way, but it is nowhere made

clear just what is its purpose.

Attention must also be drawn to what would seem to be significant omissions, especially since it is the computing side which has changed most over the past decade or so, and remarkably little discussion is included on these changes. There is only passing reference to optical discs, even though these clearly offer considerable scope for storing diagrams and half-tone pictures which are invaluable to taxonomists; these discs and their drives are now quite readily available, even if not much taxonomic data may yet have been stored on them. Taxonomists have been very slow to take advantage of these new facilities and I find little in the book which will encourage them to do so. There is no mention of optical readers or methods for the automatic incorporation of text into computer files from printed matter, and no mention of methods for scanning diagrams and pictures for the same purpose. All of these are now feasible with PCs. There is virtually no discussion of the problems of different PCs and the operating systems they use. There seems to be a tacit assumption in the book that the only PC is an IBM compatible, presumably running MS-DOS; others do exist! Although numerous software packages are mentioned, there is no indication of which PCs they will run on; coupled with limited information on their availability, this makes it almost impossible to know which, if any, of them can actually be used on any particular micro-computer system - a very real problem at the practical level!

University of Exeter R. B. Ivimey-Cook

Reference

PANKHURST, R. J. (1978), *Biological Identification. The Principles and Practice of Identification Methods in Biology*, London: Edward Arnold.

Journal of Classification 11:274-276 (1994)

R. Reyment and K. G. Jöreskog, *Applied Factor Analysis in the Natural Sciences,* Cambridge: Cambridge University Press, 1993, pp. xii + 371.

A reviewer whose own proof reading has, in the past, reduced Pearson to Peason and Bartlett to Barlett, couldn't help but warm to a book which refers to 'Applied multivariate statisitics' (sic) in the publisher's blurb on the back cover! Unfortunately this high level of *a priori* empathy was, at times, difficult to maintain once I opened the book to delve into its contents.

Applied Factor Analysis arises from the authors' earlier work *Geological Factor Analysis* published in 1976 which has now been out of print for several years. Throughout the current book the examples discussed echo the nature of the earlier work, being mainly concerned with mineralogical data. Many of these are fascinating but do perhaps limit the appeal of the book for a more general audience interested in factor analysis.

After a brief opening chapter which describes a number of applications of factor analysis reported in the geological literature, the second chapter spends 55 pages describing the basic mathematics behind factor analysis techniques. Vectors, matrices, eigenvalues, eigenvectors, etc., are covered in some detail and the account is generally clear and well laid out. But one might question whether such a lengthy description is either necessary or helpful in a book of this kind. The majority of potential readers is likely to be primarily concerned with the application of, and interpretation of results from, the methods described, rather than the fine detail of the mathematics behind them.

Chapter three gives a good account of what is termed the 'true' factor model (the common factor model), and a useful comparison of the model with principal components analysis. Chapter four gives details of fitting both the 'true' factor model and the principal components model, and spends some time on the question 'how many factors?' All the common approaches to answering this question are discussed and also included is an account of Krzanowski's cross-validatory procedure (Krzanowski 1987). Chapter four also contains a section on the principal components analysis of compositional data which draws heavily on the work of Aitchison (1986).

Reviewer's Address: B. S. Everitt, Department of Biostatistics and Computing, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF, England.

Chapter five discusses so-called 'Q-mode' methods which are concerned primarily with the classification of the objects in the sample. The authors concentrate on principal coordinates analysis (Gower 1966), and make no reference to the literature of cluster analysis. Because of this omission Chapter five is less than satisfactory, and the examples given are not convincing. Tucked away at the end of this chapter is a section entitled 'The Analysis of Asymmetry,' which deals with methods for representing graphically, proximity matrices which are not symmetric, for example, citations of journal A by journal B are not necessarily the same as citations of B by A. The method discussed is that proposed by Gower (1977) in which symmetric and skew-symmetric parts of the matrix are analyzed separately. A small example is given, which I have to confess led me to doubt the practical usefulness of the procedure.

Chapter six is concerned with correspondence analysis and the biplot. Readers are advised not to be afraid of the 'horseshoe effect.' I suspect that few readers will be! Rotation and the estimation of factor scores occupy Chapter seven and Chapter eight deals with a number of examples and case histories, followed by an account of the method of principal warps. Not to be confused with the antics of Captain Kirk and his crew, the latter is a theory based on the mathematical description of the deformation of a thin metal plate. Essentially the method deals with the analysis of shapes.

The last section of the book gives listings of a set of programs written in a language called MATLAB. The programs are available from their author Dr. L. Marcus. Personally, I doubt that 60 pages of such listings add greatly to a book of this sort, but others may disagree.

Overally there are several pluses to the book by Reyment and Jöreskog. Their accounts of factor analysis and principal components analysis and the differences between them are very good, and correspondence analysis is well described. But reducing the length of the book by cutting down on the detail in Chapter two and eliminating the program listings would have made it more readable. Nevertheless even at its present length it may appeal to some statisticians (or statisiticians!), particularly those that deal with the type of data with which the text is most concerned.

Institute of Psychiatry, London B. S. Everitt

References

AITCHISON, J. (1986), *The Statistical Analysis of Compositional Data*, London: Chapman and Hall.

GOWER, J. C. (1966), "Some Distance Properties of Latent Root and Vector Methods Used in Multivariate Analysis," *Biometrika, 53*, 325-338.

GOWER, J. C. (1977), "The Analysis of Asymmetry and Orthogonality," in *Recent Developments in Statistics*, Ed., J. Barra, Amsterdam: North-Holland.

KRZANOWSKI, W. J. (1987), "Cross-Validation in Principal Component Analyses," *Biometrics, 43*, 575-584.

Journal of Classification 11:277-279 (1994)

A. Okabe, B. Boots, and K. Sugihara, *Spatial Tesselations: Concepts and Applications of Voronoi Diagrams,* Wiley, 1992, pp. 532.

In this age of scientific specialization, ideas with broad interdisciplinary application tend often to be reinvented. Such is certainly the case with the Voronoi diagram, which despite its apparent discovery as early as the 1630s, has repeatedly appeared as a supposedly fresh idea in dozens of disciplines. Its usefulness is attested to by the huge literature on the subject, a great deal of which has been written in the last few decades. In spite of this (or perhaps because of it!), there has been little effort, beyond a number of journal review papers, to describe systematically the Voronoi diagram. Such a book is certainly needed, given the great deal that is now known about Voronoi diagrams, their cousins, the Delaunay triangulations, the many generalizations of these two concepts, and the plethora of algorithms for their computation. *Spatial Tessellations* is just such a book. In it, Okabe, Boots, and Sugihara have made a valiant attempt to draw together the many threads of this subject into a unified, broad, and yet detailed treatment.

The Voronoi diagram is easy to describe. Imagine a finite set of sites (points) in the plane. For each site define its *Voronoi region* to be the set of all points in the plane that are as close or closer to that site than they are to any other site. The collection of all such Voronoi regions is the Voronoi diagram for the given collection of sites. The boundaries of the Voronoi regions turn out to be polygons whose edges are segments of the perpendicular bisectors of the line segments joining pairs of sites. For a given pair of sites, the perpendicular bisector of the segment joining them may or may not be represented by an edge of a Voronoi region; if it is, and if we join each such pair of sites, we end up with a triangulation of the convex hull of the sites, the so-called *Delaunay triangulation.* The Delaunay triangulation can also be characterized in this way: each triple of sites (generically) determines a circle in the plane that passes through all three of them. If one of these circles is drawn, it may or may not have other sites inside it. Consider each circle that has no other sites inside it, and join the corresponding triple of sites in a triangle. This is again the Delaunay triangulation.

Reviewer's Address: Allan R. Wilks, AT&T Bell Laboratories, 600 Mountain Avenue, Room 2C-283, Murray Hill, NJ 07974, USA.

There are many beautiful properties of the Voronoi diagram and the Delaunay triangulation. For example, among all triangulations of the convex hull of the sites, the Delaunay triangulation is the one with the "nicest" triangles, in the sense that the smallest angle of any of the triangles is larger than the smallest angle in any other triangulation. The Voronoi diagram and the Delaunay triangulation may be generalized to higher dimensions, and to sites that are not just points but sets of points, such as line segments, curves, etc. Many other generalizations are possible as well. One may investigate random Voronoi diagrams, in which the sites come from some stochastic process, such as a Poisson point process, and this study leads to a wealth of further interesting properties.

To be practical, Voronoi diagrams and Delaunay triangulations must be computed, and much has been written on this subject. In fact, the Voronoi diagram is one of the unifying themes in the recent area of computational geometry.

All of these subjects, and many more, are treated in *Spatial Tessellations*. Much work has gone into the organization and selection of topics from this huge subject, and the clean result shows it. The authors begin with an historical introduction, followed by some mathematical preliminaries. There are about 50 pages of such preliminaries, most of which is probably not necessary — the book is clearly intended as a reference work so that its users are unlikely to be learning from it.

In the following two chapters, the authors define and give properties of Voronoi diagrams, Delaunay triangulations, and the many generalizations of both of them. The breadth here is quite impressive. The next chapter on algorithms shows that the authors have had personal experience in "rolling up their sleeves" and actually implementing a number of the techniques in the literature and even in generating some of their own contributions. This is most important for credibility; one only wishes that their algorithm descriptions were a little more detailed.

The remainder of the book is a collection of applications, including spatial interpolation and analysis of random point processes. Since the applications for Voronoi diagrams are extensive, this part of the book is necessarily less comprehensive, but its lack in that area is more than compensated by the depth with which their chosen subjects are treated. The book concludes with an extensive, though unannotated bibliography of several hundred references. In a bibliography of this size, some annotation would have been very welcome.

Overall, *Spatial Tessellations* stands out as an impressive overview of the area, and a most useful compendium of applications. Though the mathematical sophistication required to use the book is not small, one will nevertheless be amply rewarded for the effort needed. This book is likely to

remain an important reference for scientists in many disciplines for a number of years to come.

AT&T Bell Laboratories A. R. Wilks

Journal of Classification 11:280-282 (1994)

I. Pitas, Ed., *Parallel Algorithms for Digital Image Processing, Computer Vision and Neural Networks,* Chichester: John Wiley & Sons, 1993, pp. 395.

This book arose from a European Community-funded *ESPRIT Parallel Computing Action* in the area of *Signal Processing and Neural Networks.* It has many authors, but as the group of authors had worked together in the ESPRIT project, there is a high degree of integration of the chapters. It reads as a carefully crafted whole, much more so than most multi-authored surveys, but its scope is rather limited by the areas covered in the project.

One of the aims stated in the Preface is to provide a tutorial in the area of the title, and I felt that this goal had been met very successfully. Little knowledge of image processing, computer vision or neural networks is assumed, and essentially none on parallelization of algorithms. Indeed, the parts on image processing and vision provide a tutorial on those areas even to someone with no interest in parallel algorithms. On the other hand, the volume concludes with details of complete parallel vision systems for those interested in the nitty-gritty of implementations.

The chapter titles indicate the coverage well:

1. Introduction to parallel digital image processing
2. Low level parallel image processing
3. Parallel FFT-like transform algorithms on transputers
4. Parallel edge-detection and related algorithms
5. Parallel segmentation algorithms
6. MIMD and SIMD parallel range data segmentation
7. Parallel stereo and motion estimation
8. Parallel implementations of the backpropagation learning algorithm based on network topology
9. Parallel neural computation based on algebraic partitioning
10. Parallel neural computation based on network duplicating
11. PARALLEL EIKONA: A parallel digital image processing package
12. Parallel architectures and algorithms for real time computer vision

but also illustrate the major drawback, that the work considered is rather classical, even old-fashioned within the fields of application.

Reviewer's Address: B. D. Ripley, Department of Statistics, University of Oxford, 1 South Parks Road, Oxford OX1 3TG, UK.

There are few examples of the results of the parallel algorithms, so the reader has no idea of the overall benefits of the parallel technology. Although the design of parallel algorithms is of interest in its own right, I was left uncertain as to whether networks of transputers (the most commonly cited hardware) really did significantly outperform a single processor. The little evidence that this volume does give suggests that for whole tasks the speedup from parallelism is modest; for example, the time of 5.6 secs quoted for a FFT on 2^{18} points using 64 transputers is typical of current entry-level workstations (for example a Sun ELC).

For readers who want to see more mainstream work in this subject, Russ (1992) and Ballard and Brown (1982) provide elementary accounts with illustrations of image processing and classical computer vision respectively, and Blake and Yuille (1992) provide a modern view of computer vision, including parallel implementations of computer vision tasks not even mentioned in the current volume.

The chapters on neural networks are less satisfactory than the others. This is a large area which is covered in a rather cursory way in three chapters divided by the method of parallelization rather than by the neural network methodology. Neural networks is a field which is still developing rapidly, and there is danger in parallelizing algorithms which are unproven and rapidly superseded. The most useful parallel algorithm appears (but not from this book) to be to run a number of related models, for example nets with different starting points and/or different numbers of parameters and to combine the results. (This can be done by anyone with access to a network of machines, for example a teaching cluster during idle hours.) As deeper introductions to the area I would recommend Hertz, Krogh and Palmer (1991), Gallant (1993) and (of course) Ripley (1993).

The principal audience for this book is those working on parallel processing; it provides a useful introduction to the classical areas of its title, but this is limited by absence of the most recent problems and algorithms (especially in computer vision). For those who like your reviewer have toyed with the idea of using parallel machines in classification the message is none too encouraging; a lot of careful work seems needed for modest gains.

University of Oxford B. D. Ripley

References

BALLARD, D. H., and BROWN, C. M. (1982), *Computer Vision*, Englewood Cliffs, NJ: Prentice-Hall.

BLAKE, A., and YUILLE, A., Eds. (1992), *Active Vision*, Cambridge, MA: The MIT Press.

GALLANT, S. L. (1993), *Neural Network Learning and Expert Systems*, Cambridge, MA: The MIT Press.

HERTZ, J., KROGH, A., and PALMER, R. G. (1991) *Introduction to the Theory of Neural Computation*, Reading, MA: Addison-Wesley.

RIPLEY, B. D. (1993) "Statistical Aspects of Neural Networks," in *Networks and Chaos — Statistical and Probabilistic Aspects*, Eds., O. E. Barndorff-Nielsen, D. R. Cox, J. L. Jensen, and W. S. Kendall, London: Chapman & Hall.

RUSS, J. C. (1992), *The Image Processing Handbook*, Boca Raton, FL: CRC Press.

Journal of Classification 11:283-285 (1994)

I.A. van der Lans, *Nonlinear Multivariate Analysis for Multivariate Preference Analysis*, 1992, Leiden University, The Netherlands: DSWO Press, pp. 250.

This book examines multiattribute preference models, especially conjoint analysis and expectancy-value models, from a nonlinear multivariate data analysis framework. Its major contributions concern the various types of restrictions placed upon estimated optimal scale values, the concept of the nested vector model and principal registers, and the interesting set of procedures for implementing these models which are based on alternating least squares, iterative majorization, and unique applications of Dykstra's (1983) cyclic projection algorithm for locating optimal elements of a finite number of closed convex cones.

Chapter 1 presents a brief introduction to multiattribute preference models including a historical perspective, descriptions of such types of data, and the various uses of such data. The author appears to stress various types of Marketing applications here and throughout the monograph. Chapter 2 begins with an introduction to Nonlinear Multivariate Data Analysis (NMVA). Some of the traditional terminology is redefined in terms of new vocabulary to clarify previous "misuses" of these terms. In addition, various historical NMVA models such as MONANOVA and MORALS/MULTIPALS are briefly summarized in a conjoint analysis setting. Chapter 3 introduces compositional, self-explicated, or expectancy-value models which are described in terms of the various types of restrictions on attribute position values. The author redefines "hybrid multiattribute models" to include all models incorporating both data on attribute position values and data on direct preference scale values, and discusses more elaborate models. Finally, an algorithm is provided to obtain weighted least-squares estimates of the various parameters from such models given these restrictions. The algorithm, described in more detail in Appendices B and C, incorporates alternating least-squares, iterative majorization, and cyclic projection in a rather novel manner. Chapter 4 presents various examples of these types of restrictions. In addition, such diverse topics as to whether attribute position values should be summed or averaged across attributes, graded paired comparisons, and

---

Reviewer's Address: Wayne S. DeSarbo, Marketing & Statistics Departments, School of Business, University of Michigan, Ann Arbor, MI 48109-1234.

problems concerning solution identification are discussed. Chapter 5 presents the results of a reasonable Monte Carlo test of algorithm performance demonstrating that there are problems with local optima, but that these problems can often be minimized with multiple runs involving different starting values. Also, the first extensive empirical example is given where full-profile conjoint data on preferences for student apartments from Green and Schaffer (1991) is analyzed via the proposed procedure where dramatic individual differences in preference formation are observed. Chapter 6 begins the "second part" of the monograph which primarily deals with a variety of models for modeling individual differences. Here, an extension of Tucker's (1960) vector model called the nested vector model is proposed to model individual differences in direct and indirect preference scale values. External, internal, and intermediate analyses are discussed in a hybrid multiattribute framework. Chapter 7 begins with a discussion of inter-subject comparability. The concept of principal registers is presented as a compromise in modeling individual differences in scale values in much the same manner as principal components is utilized to model individual differences in Tucker's (1960) vector model. The main idea behind principal registers is that the "quantifications of the judgments of an individual are an idiosyncratically weighted sum of quantifications that satisfy scaling level restrictions across individuals" (p. 250). Chapter 8 revisits the apartment conjoint example presented earlier for a subset of only 25 subjects, and the nested vector model and principal registers model are illustrated in a series of analyses. Finally, Chapter 9 discusses some ideas for future research, as well as several linkages of the proposed methods with other existing methods.

The book presents an interesting framework for examining multiattribute models from an NMVA perspective. Several interesting ideas are presented for application to traditional conjoint and expectancy-value type models, especially in a Marketing context. A competent exposition of a new algorithm and its properties are provided. However, there are also a few shortcomings. One, the treatment of related techniques is very restricted and not well-developed. In particular, there are few comparisons provided in the empirical section with competing methods concerning a student apartment example. As such, it is not entirely clear what the additional complexity buys the user in terms of comparative validation rates in prediction for hold-out samples. Two, the text is rather jargony and difficult to follow in places. The redefinition of conventional terms and concepts is somewhat unnecessary and confusing. The reader is assumed to be familiar with a variety of abbreviated procedures, as very little background of many of these nor-so-well known methods is provided. Finally, the book could have been better focused in addressing a smaller subset of the wide variety of topics too lightly discussed in the monograph, and could have been reduced in length as a result.

Nonetheless, in balance, I recommend the book for researchers interested in this area.

University of Michigan Wayne S. DeSarbo

References

DYKSTRA, R. L. (1983), "An Algorithm for Restricted Least Squares Regression," *Journal of the American Statistical Association, 78*, 837-842.

GREEN, P. E., and SCHAFFER, C. M. (1991), "Importance Weight Effects on Self-Explicated Preference Models: Some Empirical Findings," *Advances in Consumer Research, 18*, 234-251.

TUCKER, L. R. (1960), "Intra-individual and Inter-individual Multidimensionality," in *Psychological Scaling: Theory and Applications*, Eds., H. Gulliksen and S. Messick, New York: Wiley, 155-167.

Journal of Classification 11:286-288 (1994)

J. D. Jobson, *Applied Multivariate Data Analysis, Volume II: Categorical and Multivariate Methods,* New York: Springer-Verlag, 1992, pp. 731.

The first volume of this book includes material on bivariate analysis for quantitative random variables, multiple linear regression, analysis of variance and experimental design. The volume under review has a more justified claim to the word 'multivariate'. The author's stated intention is to provide a second-level text for use not only by students majoring in statistics, but also by graduate students in business and the social and biological sciences. This aim seemed to me to be quite ambitious for the latter group of students, as considerable facility with matrices and linear algebra is assumed; however, relevant introductory material is summarized in an appendix.

In assessing the value of a new text book, a reviewer can ask if the material in the book is incorrect or potentially misleading, or if the book's coverage is incomplete. I have criticisms under each of these headings.

The book contains five main chapters, numbered 6-10. Chapter 6 on contingency tables discusses sampling models, the loglinear and logit models, and weighted least-squares methodology. A major criticism I have is that the word 'density' is used to refer to not only the probability function of a discrete random variable but also the relative frequency of a value in a sample. This blurring of the distinction between parameters and random variables extends to symbols used in the presentation of statistical models, which seems likely to confuse novices in the subject. Also likely to cause confusion is a statement on page 19 about the sum of multinomial random variables, and an incorrect assertion about maximum likelihood estimation of a Poisson parameter appears on page 23.

The next few chapters are better. Chapter 7 describes inference based on the multivariate normal distribution, repeated measures, multivariate regression and canonical correlation. Comments are made on outlier detection, robust estimation, influence and cross-validation, the latter topics being only briefly discussed. Chapter 8 describes multivariate analysis of variance, discriminant analysis and qualitative response models, and includes material on logistic regression and the multinomial logit model. Only rather brief comments are made on error rate estimation in discriminant analysis (for

---

Reviewer's Address: A. D. Gordon, Mathematical Institute, University of St Andrews, North Haugh, St Andrews KY16 9SS, Scotland.

which, see McLachlan 1992, Ch. 10), and I felt that some of the methodology for analyzing compositional data described by Aitchison (1986) could usefully have been included.

Chapter 9 covers principal components analysis, factor analysis, the biplot and correspondence analysis. The author discusses the use of principal components in multiple regression, illustrating the danger of discarding higher components. The factor analysis section seemed rather dated, and it was disappointing not to see material on structural equation modeling (e.g., Jöreskog and Sörbom 1988, Bollen 1989). Links between correspondence analysis and loglinear models were mentioned, but could profitably have been explained more fully (e.g., van der Heijden, de Falguerolles and de Leeuw 1989).

Chapter 10 describes cluster analysis and multidimensional scaling. The discussion of measures of similarity and dissimilarity includes on page 491 an incomplete explanation of the relationship between Mahalanobis and Euclidean distances, which could confuse unwary readers. The description of clustering procedures concentrates on several standard criteria obtainable within Lance and Williams's (1966) general agglomerative algorithm. On page 516, the author defines ultrametricity in terms of the absence of inversions in a dendrogram, and gives an incomplete set of conditions necessary for this property to hold. In a discussion of the distorting effects of clustering criteria on page 524, the single link criterion is said to be space conserving (instead of 'contracting') and the complete link criterion is said to be space diluting (sic); this is one of the more amusing examples of inattentive proofreading. On succeeding pages, it is asserted that the centroid method is space contracting and Ward's method is space dil[a]ting, whereas these two criteria can also be space conserving (Ohsumi and Nakamura 1989). The cluster analysis section also includes a description of some test statistics for determining the number of clusters in the data, and an outline of simulation studies conducted by G. W. Milligan and colleagues (e.g., Milligan and Cooper 1985, 1988). The multidimensional scaling section concentrates on classical (metric) MDS and nonmetric MDS. The example chosen to illustrate classical MDS is unfortunate: inter-city flight times can be expected to be asymmetric (and correct specification of the locations of two cities does not suffice to recover the map: page 583).

As far as coverage of material is concerned, each individual can be expected to have slightly different views of what it is appropriate to include. I have already suggested a few possible additions, but readily admit the subjective nature of these views. However, I would argue more strongly that a text on multivariate data analysis should include discussions of the treatment of incomplete data (e.g., Little and Rubin 1987).

Thus far, I have concentrated on assessing the presentation of the theory. Despite the comments made on Chapter 10, a generally strong feature of the book is the analysis and discussion of examples, for which the author reports using the BMDP, SAS, SPSSX and CLUSTAN packages. Each chapter is followed by a set of practical exercises and some questions of a theoretical nature. There is also an appendix that presents 22 data sets, which are available on a diskette.

It is possible that a revised edition of this book could be of value to its intended audience, but I do not feel able to recommend the current version.

University of St Andrews A. D. Gordon

References

AITCHISON, J. (1986), *The Statistical Analysis of Compositional Data*, London: Chapman & Hall.

BOLLEN, K. A. (1989), *Structural Equations with Latent Variables*, New York: Wiley.

JÖRESKOG, K. G., and SÖRBOM, D. (1988), *LISREL 7, A Guide to the Program and Applications*, Chicago: SPSS.

LANCE, G. N., and WILLIAMS, W. T. (1966), "A Generalized Sorting Strategy for Computer Classifications," *Nature, 212*, 218.

LITTLE, R. J. A., and RUBIN, D. B. (1987), *Statistical Analysis with Missing Data*, New York: Wiley.

MCLACHLAN, G. J. (1992), *Discriminant Analysis and Statistical Pattern Recognition*, New York: Wiley.

MILLIGAN, G. W., and COOPER, M. C. (1985), "An Examination of Procedures for Determining the Number of Clusters in a Data Set," *Psychometrika, 50*, 159-179.

MILLIGAN, G. W., and COOPER, M. C. (1988), "A Study of Standardization of Variables in Cluster Analysis," *Journal of Classification, 5*, 181-204.

OHSUMI, N., and NAKAMURA, N. (1989), "Space-Distorting Properties in Agglomerative Hierarchical Clustering Algorithms and a Simplified Method for Combinatorial Method," in *Data Analysis, Learning Symbolic and Numeric Knowledge*, Ed., E. Diday, New York: Nova Science Publishers, 103-108.

VAN DER HEIJDEN, P. G. M., DE FALGUEROLLES, A., and DE LEEUW, J. (1989), "A Combined Approach to Contingency Table Analysis Using Correspondence Analysis and Log-Linear Analysis (with Discussion)," *Applied Statistics, 38*, 249-292.

Journal of Classification 11:289-290 (1994)

J.-P. Benzécri, *Correspondence Analysis Handbook*, (Translated by T. K. Gopalan), New York: Marcel-Dekker, Inc., 1992, pp. 665.

This is a translation of Benzécri's 'bible' of correspondence analysis, previously available only in French. Professor Benzécri is something of a guru of the French school of data analysis, which has, at least in the past, depended heavily on the use of the single technique of correspondence analysis. In contrast, in other countries, a more eclectic approach has been taken, using a broad spectrum of different kinds of statistical tools. Over the last ten years or so, however, researchers have explored the relationships between correspondence analysis and other methods and have attempted to identify the sorts of problems that each method is best suited to tackle.

If correspondence analysis is but a single technique, it is related to many others, and the discussion in this book does show some of these relationships. It illustrates the artificial nature of the division of statistics into sets of different tools with different names. It also shows how different methods may be regarded as special cases or slight variants of other methods, though it does all this from the partisan perspective of a proponent of correspondence analysis.

With its many examples of correspondence analysis being applied (in different ways) the book provides an excellent illustration of how sensitive and sophisticated use of a single technique can shed light on data in many different ways. It serves to support the position that a thorough grasp of a few techniques is better than a weak grasp of many.

Having said all that, the book is in some ways (perhaps inevitably) rather old-fashioned: it does not, for example, relate the method to very similar or identical techniques developed by other authors and which have inevitably been accorded different names (such as dual scaling). Nor does it relate the method to log-linear modeling, a tool which is widely used to tackle categorical data problems, as is correspondence analysis.

The translation is sound, though there are one or two places where the English is stylistically rather awkward and some relaxing of the literal translation from the French would have been beneficial. The approach is slow and detailed.

Reviewer's Address: D. J. Hand, Statistics Department, The Open University. Walton Hall, Milton Keynes MK7 6AA, England.

I was not impressed by the typography and layout which I found rather ugly and difficult to read. However, I found only a few typographical errors.

So much for the overall nature of the book. Now what about its content?

The volume is divided into five parts, each commencing with a page or two summarizing the contents of each of the chapters in that part.

Part I, on the theory of correspondence analysis, contains 12 chapters, outlining the geometrical language and basic mathematics. In general the book does not require a deep mathematical background, and this part includes chapters on vector and affine geometry and on Euclidean geometry. Part II works though (by hand) a simple example to illustrate the steps of an analysis. Part III introduces correspondence analysis by computer, and includes a complete Fortran program for correspondence analysis. Part IV describes sixteen examples to illustrate the range and diversity of applications of correspondence analysis. Part V of the book is rather anomalous, and it is not clear to me that it was appropriate to include it. This part is concerned with agglomerative hierarchical clustering, which is described as 'a companion method to correspondence analysis'.

There are two other structural peculiarities which are worth mentioning. First, although the contents list shows there to be an index, in fact there is none, and what is called an index is really a glossary, though with rather inadequate definitions. Instead of giving page numbers, this index/glossary gives section numbers. Secondly, there appears to be no list of references, instead giving sufficient details of the references in the text where they appear. This is not ideal.

In summary, then, we have a translation of a book promoting correspondence analysis, rather unusual in style, but which nevertheless provides a good introduction to the technique (both in practice and in theory). For someone new to the method this would be a good — and accessible — place to start, though it is not short.

The Open University D. J. Hand

Journal of Classification 11:291-294 (1994)

R. Steyer, K. F. Wender, K. F. Widaman, Eds., *Psychometric Methodology, Proceedings of the 7th European Meeting of the Psychometric Society in Trier, July 29-31, 1991,* Stuttgart and New York: Gustav Fischer Verlag, 1993, pp. xiv + 596.

The proceedings consist of 111 contributions on Item Response Theory, Multivariate Analysis (including Classification and Factor Analysis), Multidimensional Scaling, Psychophysics and Information Processing, Structural Equation Models and some other topics. The editors decided to arrange the papers alphabetically according to the first author's name because sometimes it was hard to specify their topics.

Item response theory (IRT) and related areas of educational and psychological measurement attracted the highest number of contributions. Different aspects of validity and reliability of tests, item/test and response biases are presented in the papers by M. Cuesta et al., J. Gómez-Benito and M. Forns-Santacana, P. Kovacevic et al., D. Hutchison and I. Schagen, R. R. Meijer and K. Sijtsma, P. Dickes and C. Reinhardt, A. D'Onofrio et al., I. P. Vasilescu, M. W. Browne, J. W. Graham et al., R. Nandakumar, F. J. Oort, C. Pèrez and J. L. Padilla, P. Westers. In the paper by E. Boekkooi-Timminga, which was written as a reaction to the work of Stocking (Stocking 1990), a method for determining the optimum sample of examinees for estimating item parameters in IRT is proposed. The Rasch model (Rasch 1960), its generalizations, applications and related topics are discussed by J. Rost and M. von Davier, W. J. van der Linden and T. J. H. M. Eggen, H. Kelderman, M. G. H. Jansen, H. Hoijtink, Á. Münnich defines and describes some probabilistic choice models within the framework of IRT. B. W. Junker and W. F. Stout investigate robustness of ability estimates when unidimensional IRT models are used with multidimensional data. I. W. Molenaar and Á. Münnich present a formula that transforms latent traits measured by cumulative items into point items and vice versa. R. Zwick determines admissible transformations of achievement scales for the National Assessment of Educational Progress (NAEP) survey. G. J. Mellenbergh describes and analyzes the unidimensional latent trait model for continuous item responses. R. Janssen et al. apply Embretson's (Embretson 1984) multicomponent latent trait model to synonym tests. I. Partchev presents the Bulgarian standardization and results

---

Reviewer's Address: V. Pliner, Medical & Health Research Association of New York City, MIT Study, 125 Worth Street, Box 44, New York, NY 10013, USA.

of Wechsler's test of intelligence (HAWIK-R 1984) for children. H. Pitariu et al. describe an expert system which they developed to assist dealing with California Psychological Inventory (CPI) (Gough 1987). G. H. Maassen and W. Akkermans introduce a probability model for sociometric status determination with rating scales. P. Blahuš discusses relationship between measurement theory and probability theory with special emphasis on definition of random variable in terms of representational measurement theory. Some aspects of psychological measurement theory are also handled by B. V. Pulkin, S. Logar-Djuric et al., M. C. Viladrich and M. D. Riba, S. A. Marinov, and some applications using psychometric tests and questionnaires are described by J. M. Plusnin and A. A. Putilov, I. L. Chepkasov.

Various techniques and aspects of multivariate analysis attracted quite a few contributions also. Two of them (by K. Schweizer and by R. Moreno and J. A. Pérez-Gil) compare different techniques for classifying variables. A. Vargha proposes using ANOVA for analyzing dependency of a dichotomous variable on a set of several discrete variables. J. M. F. ten Berge and H. A. L. Kiers offer a computational procedure for solving the minimum rank problem in factor analysis. M. Ihara and Y. Kano focus on selection of variables in factor analysis. W. P. Krijnen proposes a restricted PARAFAC (Harshman and Lundy 1984) method to represent three-way data. H. Neudecker treats a technical problem for Jöreskog's Image Factor Analysis model (Jöreskog 1962). I. A. van der Lans, W. A. van der Kloot deal with preference data analysis in their papers. M. Paechter and D. Bartram conduct a comparative analysis of different cross-validation techniques for prediction quality estimation in multiple regression. Techniques for handling missing data are discussed by S. van Buuren and J. L. A. van Rijckevorsel, P. E. Cheng. E. J. Manfred and G. J. Rezendes present an application of regression analysis to a placement procedure at the U.S. Coast Guard Academy, A. N. Sokolov applies ANOVA to a problem of psychophysics and E. van der Burg and G. B. Dijksterhuis offer an application of nonlinear redundancy analysis and of nonlinear canonical correlation analysis. Several more authors handle other models and aspects of multivariate statistical analysis, e.g., P. Verboon, V. Zdravkov and M. Boev, Y. Kano, A. J. Kutylowski, I. Rentschler et al., R. R. Macdonald.

Correspondence analysis, otherwise known by many other names, including homogeneity analysis, dual scaling, optimal scoring, scalogram analysis, is represented in the papers by M. Th. Markus and R. A. Visser, H. Matschinger and M. C. Angermeyer, J. Pannekoek. M. Kieser, P. G. M. van der Heijden and A. Mooijaart deal with other approaches to contingency tables analyses. M. A. J. van Duijn and U. Böckenholt in their contributions introduce models for the analysis of contingency tables of count data.

Papers on multidimensional scaling (MDS) and unfolding techniques cover several aspects of the topic. P. Arabie presents a brief survey of multidimensional scaling based on the city-block metric, abstracted from the 1991 Psychometric Society Presidential Address (Arabie 1991). G. Bove and F. Critchley as well as T. Saito in their papers deal with multidimensional scaling for asymmetric proximities. Several contributions (authors: Ch. A. Izmailov, A. N. Lebedev, S. Lewandowsky and D. A. Newman, K. Davydova and Ch. A. Izmailov) present MDS applications to psychophysical problems of perception. P. J. F. Groenen proposes using a modified Leverrier-Faddeev algorithm to simplify the computation of the inverse of a data matrix for partitioned block designs in MDS. The empirical study (on moral evaluation) by J. Tournois uses MDS to determine the effects of measurement constraints on the measurement results. For analysis of individual differences in MDS A. F. M. Nierop presents a modification of INDSCAL model. J. P. Sutcliffe offers his view on comparison of MDS, numerical taxonomy and logical taxonomy. Two papers, by W. H. van Schuur and by P. V. Zysno, deal with unfolding models which are often considered as a subtype of MDS.

Structural equation models, also known as path analysis, covariance structure analysis, errors-in-variables and equations models, or by more specialized names such as LISREL or confirmatory factor analysis models, are represented by several papers. P. Notz, M. Eid and R. Steyer introduce a computer program for the analysis of latent state-trait models, a special kind of structural equation models, and M. J. Schmitt and R. Steyer propose a latent state-trait model for social desirability. A. Satorra and W. E. Saris conduct power analysis of the chi-square goodness-of-fit test for simultaneous misspecification on several parameter restrictions. D. Kaplan and R. N. Wenger discuss specification error propagation problem. T. Blank and J. Reinecke compare the LISREL and LISCOMP approaches for long-wave panel analysis with categorical variables. T. Raykov describes a model that allows for studying various patterns of change in true test scores over time. In the paper by M. Eid et al. a psychological application of structural equation models is handled. L. M. Collins et al. and A. Camstra use cross-validation techniques for model selection.

Short descriptions of computerized software are provided by A. Bartkowiak on teaching basic ideas of statistics (PS-STAT), by S. Gasik on social processes description and simulation (WSL language) and by H. Klein on text analysis (INTEXT/PC).

A number of papers discuss such aspects of Psychophysics and Information Processing as visual perception, imagery and verbal processes in memory, cognitive learning processes, the effect of emotional states on memory and other cognitive processes, psychophysical analysis of decision-making and some others.

As a collection of papers on a wide variety of topics, the volume can not claim to give a comprehensive treatment of the subject of Psychometric Methodology, but can be recommended to specialists as it introduces readers to the latest developments in the covered areas and is extensively illustrated by real-world examples.

Medical & Health Research Association of New York City

V. Pliner

References

ARABIE, P. (1991), "Was Euclid an Unnecessarily Sophisticated Psychologist?" *Psychometrika, 56*, 567-587.

EMBRETSON, S. E. (1984), "A General Latent Trait Model for Response Processes," *Psychometrika, 49*, 175-186.

GOUGH, H. G. (1987), *California Psychological Inventory. Administrator's Guide,* Palo Alto, CA: Consulting Psychologist Press.

HARSHMAN, R. A., and LUNDY, M. E. (1984), "The PARAFAC Model for Three-Way Factor Analysis and Multidimensional Scaling," in *Research Methods for Multi-mode Data Analysis*, Eds., H. G. Law et al., New York: Praeger, 122-215.

HAWIK-R (1984), Hamburg-Wechsler Intelligenztest für Kinder - Revision 1983 [Hamburg-Wechsler Test of Intelligence for Children - Revision 1983: HAWIK-R], Edited and revised by U. Tewes (2nd ed.), Bern: Huber.

JÖRESKOG, K. G. (1962), "On the Statistical Treatment of Residuals in Factor Analysis," *Psychometrika, 27*, 335-345.

RASCH, G. (1960), *Probabilistic Models for Some Intelligence and Attainment Tests*, Copenhagen: Denmarks Paedogogiske Institut.

STOCKING, M. L. (1990), "Specifying Optimum Examinees for Item Parameter Estimation in Item Response Theory," *Psychometrika, 55*, 461-475.

Journal of Classification 11:295-296 (1994)

J. Scott, *Social Network Analysis - A Handbook,* London: Sage, 1991, pp. 210.

This book describes a methodology for analyzing relational data. Its aim is 'to simplify the techniques of social network analysis', 'to expand the understanding of social network analysis by motivating people to acquire the more complex knowledge as and when they need it' and 'to bridge the gap between theory and practice of network analysis'. The book concentrates on identifying the key social network concepts used in assessing network structure (e.g., density, centrality, cliques, positions, roles). The first such introductory book was written by Knoke and Kuklinski (1982). But as the field of network analysis has developed very rapidly in the last two decades, a book with these goals is very welcome.

The book consists of eight chapters and an appendix. After the introductory chapter there is a very nice discussion of the development of social network analysis, looking at its origins in the social psychology of groups and at its subsequent development in sociological and social anthropological studies of factories and communities. The chapter concentrates on the theoretical ideas which emerged in this work and shows how this was connected with the growing technical complexity of the work carried out from the 1970s. Chapter 3 discusses the organization, storage and selection of relational data. Chapter 4 introduces the basic building blocks of social networks: representation of a network as a graph of points and lines. Other concepts like distance, direction and density are also defined. In Chapter 5 centrality of points and the centralization of whole networks is presented and discussed. Chapter 6 examines some of the principal concepts proposed for the investigation of sub-groups within social networks: components, cores and cliques. In Chapter 7 the structure of positions and roles which are defined by social relations is presented. Chapter 8 looks at the approaches for displaying relational data (e.g., metric and non-metric multidimensional scaling, principal component analysis). In the Appendix there is also an introduction to and comparison of the main computer programs for social network analysis.

The risk of such an introductory book is that many matters may be oversimplified. For example, this can be noticed in the chapter dealing with positions, roles and clusters. This chapter is also less up-to-date because

---

Reviewer's Address: A. Ferligoj, Faculty of Social Sciences, Kardeljeva pl. 5, 61000 Ljubljana, Slovenia.

several important results on blockmodeling, which clarify the analysis of positions and roles, were published during the last few years, especially in the special double issue of the journal *Social Networks* (Borgatti 1992). Nevertheless, this book will be useful to teachers, students and researchers who would like to get a first insight into social network analysis.

Ljubljana A. Ferligoj

References

BORGATTI, S. P., Ed. (1992), "Special Issue on Blockmodeling," *Social Networks 14*.
KNOKE, D., and KUKLINSKI, J. H. (1982), *Network Analysis*, Beverly Hills: Sage.