

- names", "Compound surnames", "Compound surnames with prefixes", "Royalty and nobility", "Saints and ecclesiastical dignitaries", "Married women", "Change of title or status", "No-surname headings", "Foreign personal names" (Arabic, Burmese, Chinese, Indian, Indonesian, Japanese).
- 8 A phrasing, used by D.B.Cleveland and A.D.Cleveland (*Introduction to indexing and abstracting*. Littleton, CO: Libraries Unlimited 1983. p.12), which aptly characterizes N.Knight's book (note 7).

Dr.Werner Bies, Universitätsbibliothek der FU Berlin, Garystr. 39, D-1000 Berlin 33.

Bock, Hans H.: **Classification and Related Methods of Data Analysis.** Proceedings of the First Conference of the International Federation of Classification Societies (IFCS), Aachen, FRG, 29 June - 1 July 1987. Amsterdam/New York/Oxford/Tokyo: North Holland 1988. XV,749p. ISBN 0-444-70404-3

Eighty-six invited or selected and refereed papers bring together a wide range of topics and perspectives in the field of classification, clustering, and data analysis. They may be subdivided roughly into four groups. In reviewing them, we refer to the concise summary given by the editor in his preface.

The papers in the first group deal with various mathematical and statistical aspects of clustering and classification. Their scope is indicated by keywords like (dis)similarity measures, clustering algorithms, classification criteria, optimal partitions, trees and hierarchies, fuzzy classification methods, mixture analysis, discriminant analysis and image segmentation. Special emphasis has been given to topics like probabilistic clustering models, inferential statistics, the assessment and evaluation of clustering, and cluster or homogeneity tests.

The second group is devoted to consensus and aggregation methods for relational and qualitative data. The papers survey the theoretical background as well as applications in the social and biological sciences, including methods for biological taxonomy, the reconstruction of phylogenies, and the analysis of molecular structures in chemistry and microbiology.

The third group covers topics from multivariate statistics and exploratory data analysis in the broad sense. These papers deal with (ordinary and constrained) multidimensional scaling, related graphical data representation methods, correspondence and canonical analysis (comprising its nonlinear and loglinear extension), seriation methods, multiple and pairwise comparisons, ordinal network and graph analysis, algebraic methods from formal concept analysis, and model selection criteria.

A major topic of the conference was the presentation of relevant software packages and the investigation of the interrelationship and mutual impact between classification, retrieval, and expert systems. The papers of the last group reflect these activities and describe or analyze expert and data management systems developed for classification purposes.

From many topics worth mentioning, we have selected some which are important in practice but under-represented in this volume (as in many others before). The

handling of measurement errors is such a neglected topic; only two papers were found dealing directly with it. One adopts methods from the syntactic pattern recognition using a formal language approach to inexact graph matching (M.Kaul). It combines two interesting aspects, the handling of structural information, and error handling. The structural knowledge is incorporated into a graph grammar, and error handling (of possibly distorted input graphs) is performed with an error-correcting parser. The reconstruction of phylogenesis by weighted genetic distances (P.O.Degens) is the other approach tackling with the error problem. It is based on a generalization of average linkage for a Gaussian error model.

A serious problem which should be given more attention concerns the danger of losing track of things in the mass of details. Above all, this shortcoming is felt by practitioners who have to select the model and perform the analysis. Software tools with a large number of routines are available now, but less guidance is present in making decisions about the order in which different algorithms should be applied. Expert systems are in development to overcome this problem for cluster analysis (E.Backer) and exploratory (election) data analysis (F.Gebhardt). These attempts are still in their initial state. Designed for a general application, they require general principles for their foundation, not available as yet. In the future such projects will certainly promote theory-oriented studies. Approaches like the new information-theoretic measure of complexity for model selection (H.Bozdogan) may be a good starting point for this. It provides a new criterion of goodness for the fit of a model. Some values from the new criterion are compared with those obtained by the fundamentally different Akaike Information Criterion.

The volume contains a noticeable number of survey lectures and invited papers. The opinions may differ on the usefulness of this practice. The classification problem dealt with from the pattern recognition point of view (H.Niemann) stands out from them on account of its succinct formulation. A negative example is, in my opinion, the paper on the modern fundamental measurement theory (R.D.Luce). Continuing an old tradition in mathematical psychology, 'measurement' is confused with 'metrization', so that the theory appears neither modern, nor fundamental.

The contributions, although varying in their quality, convey a good overview about the state of the art in numerical classification and related methods of data analysis. Many of the papers demonstrate the usefulness of these methods in applying them to real data, e.g. from astronomy, biology, marketing, psychology, the social sciences, etc. The subject index of this volume facilitates cross referencing between these practical applications and the methods. Mathematicians, statisticians and data analysts as well as researchers and practitioners will find many stimuli for their work.

Peter Jaenecke

Dr.P.Jaenecke, Standard Elektrik Lorenz AG, Forschungszentrum Fachbereich Bürokommunikation, Ostendstr. 3, D-7530 Pforzheim