Book Reviews

JAMBU, M.; LEBEAUX, M.-O.: Cluster Analysis and Data Analysis.
Amsterdam-New York-Oxford: North-Holland 1983. XXIV,898p., ISBN 0-444-866345-5

The book consists of two parts: Cluster analysis and data analysis, although the first is often considered as a part of the second. The author defines the first part as "the set of computer procedures which build up or recognize hierarchies or partitions of underlying data sets", the second as "the set of computer procedures which describe, recognize or identify structures of underlying clouds of points, usually points of a low-dimensional space built from the data sets...". It does not become clear why the author speaks of computer programs instead of algorithms. The reason may be that three quarters of the book consist of listings and descriptions of FORTRAN programs for the different altorithms described in the theoretical part of the book.

The theoretical part starts after a brief introduction to the classification problem with a chapter on correspondence analysis and another one on properties of hierarchical classifications. A description of distance and similarity measures follows. The following chapters are devoted to classical and fast algorithms for hierarchical classifications, algorithms for determining partitions and overlapping clusters and interpretational aids and validity of classifications. Exact definitions, theorems and their proofs are given but they are always explained and justified by practical examples.

The book can be recommended to users of classification methods who want to create a program library and also to those who are interested in the theoretical background of the methods. Peter Ihm

Prof.Dr.P.Ihm, Inst.f.Medizinische Statistik,
Ernst-Giller-Str.20, D-3550 Marburg

DAHLBERG, Ingetraut (Ed.): Klassifikation als Werkzeug der Lehre und Forschung (Classification as a tool in teaching and research). Frankfurt/ Main: Indeks Verlag 1986. 184p. = Studien zur Klassifikation, Bd.16; ISBN 3-88672-015-2

In this volume of lectures under review, the attempt is made, among other things, to focus the attention of those users engaged in teaching and research on the concept of "classification".

In the Preface it is emphasized "how the methods of classification as the organisation of knowledge attend on and exercise an essential influence on the research process and, indeed, may even become a direct aid in research. First of all, concept analysis and concept definition form the basis of all research, for without a firm grasp of concepts as units of our knowledge, problems can neither be scientifically recognized as such nor solutions found for them. The next step, which usually consists in the collecting and analysis of data, demands classification knowledge of the correct assigning of data and their grouping. In a further, third step, data structuring and data ordering must be undertaken, and it is sure that without an explicit knowledge of existing classification possibilities, this can only be done in an amateurish fashion" (Dahlberg).

To consider classification under the aspect of the "ordering of knowledge" has not only become important in respect of the rationalisation of traditional methods of research. The reapid development of information technology marked by the transition from "data processing" to "knowledge processing" also makes new demands on the methods of classification. In this field, a large number of problems have, up to now, found either no solution whatsoever or have been but unsatisfactorily solved. It is not so much problems of computer technology or programming techniques which now lie to the fore, but mainly problems connected with the form of human knowledge processing, with semantics, and with suitable concept ordering.

If systems based on knowledge (expert systems) are also to carry out "intelligent" functions and be employed to solve scientific problems, then we must not forget that in doing so a great number of characteristics considered a prerequisite for intelligence must also be taken into account in a certain way when solving problems with the aid of a computer.

NEDOBITY lists these characteristics in his paper and comes to the conclusion that they are not only of great interest for research. They are also an integral part of classification activity itself. What is more, an effective man-computer interaction is surely hardly realisable if the user's conceptions of order or the method of ordering knowledge in the user's memory has not been adjusted to the order of knowledge fed into the computer system. In his contribution, Nedobity points out this psychological aspect of classification and knowledge techniques. Here, the comparison of the various approaches to the presentation of knowledge is of particular interest:

- in artificial intelligence
- in classification, and
- in cognitive psychology.

The description of the consequences and advantages of these approaches stimulates analogous reflections which may also contribute to the optimisation of relevant methods and aids. When from all approaches and conclusions it follows over and over again that knowledge is "structured" information, that "problem solving behaviour is a structured process", that "the human brain operates in symbolic pictures" (Nedobity), then in my opinion the question - would a structured picture-form of knowledge presentation as opposed to other methods not have certain advantages? - is irresistible. As is known, the structured formula picture of a chemical compound possibly represents a very great deal of knowledge in concise form and, compared to the relevant chemical name of the compound, distinguishes itself in a high degree of condensed information and transparency.

Of course, the structured form of presentation of any given store of knowledge pre-supposes a certain formalisation of existing knowledge in the natural language analogous to the structured formula picture in chemistry, a degree of formalisation adapted to the relevant purpose of application, as has been done, for example, in a subject thesaurus. From these points of

https://doi.org/10.5771/0943-7444-1988-2-99-1 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

view, knowledge presentation would then mean a thesaurus-like presentation of knowledge structure, and the processing of knowledge the reshaping of knowledge structures.

UNGVARY's contribution, "The use of the thesaurus method in knowledge conveyance", runs on similar lines. Ungvary, however, confines himself primarily to the theoretic aspects of the use of the thesaurus method in knowledge "conveyance" and does not go into the problem that, for example, good use may also be made of thesaurus changes in the acquisition of knowledge. A discussion of connexions to other possibilities of applying the thesaurus principle, for example in prognosis (Dobrow) or problem solving (Bauer), would have been useful.

The wide thematic spectrum of the contributions in this volume with all the questions that are touched on, but still open is, at the same time, an appeal to promote classification research even more than has been the case up to now.

"But classification training itself should be improved and coordinated in all areas; above all, essential fields of classification science should be offered to students of all disciplines at colleges, universities and training centres as a subject which not only teaches methods of order, but also gives a very necessary overall view of the various fields of knowledge, which draws the individual out of the isolation inherent in a specialised field of knowledge, and which gives him again insight into the general view of knowledge" (Dahlberg).

In this review it was only possible to go into a few of the particularly acute problems and focal points taken from the overall theme of the volume. As already stated in the Preface, the problems presented here have, for the meoment, only been laid open for discussion. But one thing is sure: this volume should be stimulating for all specialists, well beyond those who belong to the circle members of the classification society.

Gerd Bauer

Dr.rer.nat.Dr.sc.phil.G.Bauer
Am Treptower Park 50, GDR-1193 Berlin

MANIEZ, Jacques: **Les langages documentaires et classificatoires: conception, construction et utilisation dans les systemes documentaires.**
(Documentary and classification languages: Their conception, construction and use in documentary systems).
Paris: Les Editions d'Organisation 1987. 291p. ISBN 2-7081-0833-6.

Jacques Maniez's book and that of G.Van Slype (Indexation Languages: Design, Construction and Use in Documentary Systems) published in the same year and by the same publisher, are a joint project dealing with documentary languages in general, and principally intended for professionals, though of possible interest to readers interested in information techniques and methods of content analysis.

The major portion of J.Maniez's book thus deals with classificatory languages. The other portion discusses documentary languages, which are presented in the light of linguistic theory rather than from a practical and descriptive point of view. The advantages and drawbacks of documentary languages are related to the properties of natural language. J.Maniez uses his competencies as a linguist to provide an in-depth analysis of the principal relationships used in documentary languages (synonymy, hierarchy and association) as well as the operations of characterization and condensation of a document and research in automated documentation.

In the principal portion of J.Maniez's work, the first chapter gives us a definition of important terms and presents a typology of classification systems. In the second chapter, the author shows us how to use these tools and describes those which are the most well-known (Dewey Decimal and Universal Decimal Classification, Colon Classification, Library of Congress Classification and the Bliss Classification). Finally, an interesting chapter is devoted to the construction and maintenance of these classificatory languages.

As shown in the above description of these different chapters, J.Maniez's book is of interest at different levels:

(1) The practicing classificationist whose knowledge was principally acquired through practical experience will find in this book a theoretical presentation which will allow him to better understand the basic principles of the design and use of classification systems. Certain empirical rules which he discovered through trial and error will here find their justification - if not theoretical, then at least within the framework of a coherent set of practical considerations. On numerous occasions, the author's linguistic background enables him to explain and criticize certain rules in current use. For example, he presents the notion of facet by comparing it to case in a language, thus allowing for a better understanding of the concept, as well as the advantages and drawbacks of its use. J.Maniez does not hesitate to present projects which are still in a research stage, such as G.Salton's famous SMART system. The author' s presentation of the foundation of automated classification techniques is in itself a lesson in pedagogy.

(2) The second level will be useful to those who, having little knowledge of classification systems, will find the essential elements needed for an understanding of this field. The main classification systems are presented in a sufficiently precise manner to allow the reader to understand their interest and their functions. Incidentally, the classificatory techniques are covered with true pedagogical talent by the author. Each classification system is presented in a highly structured form which successively describes the principles, the working instruments, the use of these instruments by the classifier and by the reader, as well as a critical conclusion. Extracts of tables and indexes as well as precise examples provide a very concrete and practical understanding of these tools.

(3) The last level, to be found in chapter III, will be extremely valuable to those who must design and set up a classification system in a library. This chapter surely sets this book apart from others that are limited only to explanation or description. Placing himself in a familiar context which corresponds to a specialized library of approximately 6000 works, J.Maniez constructs a veritable case study which allows us to determine the characteristics of two classification systems which can answer the needs of this case (the final choice is left to the head of the establishment).

https://doi.org/10.5771/0943-7444-1988-2-99-1 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY