GENERAL INTRODUCTION: PURPOSE, UNDERLYING IDEAS, AND SCOPE OF THE BOOK

by

I. Van Mechelen
R. S. Michalski

Categories and Concepts Theoretical Views and Inductive Data Analysis,
I. Van Mechelen, J. Hampton, R. S. Michalski and P. Theuns (Eds.), pp. 1-7,
Academic Press Ltd., 1993.

Categories and Concepts

Theoretical Views and Inductive Data Analysis

edited by

Iven Van Mechelen
Katholieke Universiteit Leuven
Belgium

James Hampton
City University London
UK

Ryszard S. Michalski
George Mason University
USA

Peter Theuns
Vrije Universiteit Brussel
Belgium

ACADEMIC PRESS
Harcourt Brace Jovanovich, Publishers
London San Diego New York Boston
Sydney Tokyo Toronto

Cognitive Science Series

GENERAL EDITORS

Barry Richards *Imperial College, London*
Keith Stenning *University of Edinburgh*

ADVISORY BOARD

Michael Arbib *University of Southern California*
Keith Brown *University of Essex*
Dov Gabbay *Imperial College, London*
Barbara Grosz *Harvard University*
Franz Guenthner *University of Tubingen*
Jaakko Hintikka *Florida State University*
Ewan Klein *University of Edinburgh*
Philip Johnson-Laird *Princeton University*
Gareth Matthews *University of Massachusetts (Amherst)*
Brendan McGonigle *University of Edinburgh*
Terry Myers *University of Edinburgh*
Stellan Ohllsson *University of Pittsburgh*
Stanley Peters *Stanford University*
Mark Steedman *University of Edinburgh*
Henry Thompson *University of Edinburgh*
Yorick Wilks *New Mexico State University*
Richard Young *MRC Applied Psychology Unit (Cambridge)*

Collections

1. Reasoning and Discourse Process *T. Myers, K. Brown and B. McGonigle (eds), 1986*
2. New Directions in Semantics *E. Le Pore (ed.), 1987*
3. Language Perception and Production: Relationships between listening, speaking, reading and writing *A. Allport, D. G. MacKay, W. Prinz and E. Scheerer (eds), 1987*
4. Current Research in Natural Language Generation *R. Dale, C. Mellish and M. Zock (eds), 1990*
5. Representing Musical Structure *P. Howell, R. West and I. Cross (eds), 1991*
6. Algorithms, Abstraction and Implementation: Levels of Detail in Cognitive Science *C. Foster, 1992*

Monographs

Agreement and Anaphora *P. Bosch, 1983*
Temporal Representation and Inference *B. Richards, I. Bethke, J. van der Does and J. Oberlander, 1989*

Contents

1

General Introduction: Purpose, Underlying Ideas, and Scope of the Book

IVEN VAN MECHELEN
Departement Psychologie, Katholieke Universiteit Leuven, Tiensestraat 102, B–3000 Leuven, Belgium

and

RYSZARD S. MICHALSKI
Center for Artificial Intelligence, George Mason University, Fairfax, VA 22030–4444, USA

1 Categories and concepts

For humans facing a host of experiences in a complex world, an important activity is to break up these experiences into meaningful, more manageable components. This is the basic problem of categorization. Categorization begins at the early stages of any sensory perception when incoming stimuli are linked to more abstract categories, as appears from the fact that equal-sized physical differences between stimuli are perceived as larger or smaller depending on whether they are in the same category or in different ones (Harnad, 1987). Furthermore, the acquisition of concepts and categorical structures is a fundamental part of any learning process. Categories or their mental representations are pre-eminent objects of thinking processes. Also, in any language a large portion of words can be considered to consist of category labels; actually, a significant part of any human communication consists of an exchange of categorical information.

Categorization not only plays a central role in everyday life, but is also a research topic for many domains of science. Well known, for example, are categorization efforts in biology, in particular the development of

CATEGORIES AND CONCEPTS: THEORETICAL VIEWS AND INDUCTIVE DATA ANALYSIS
ISBN 0-12-714175-8
Copyright © 1993 Academic Press Ltd
All rights of reproduction in any form reserved

taxonomic systems for fauna and flora. In archaeology and anthropology, the retrieval of categorical structures of artefacts and of items of cultural worth has always been a central task. Categorization also occupies a prominent place in clinical domains, such as medicine and psychiatry, where therapeutic interventions are preceded by a diagnostic categorization of the problem. Such categorizations are typically a reflection of the underlying causal dependencies.

This book concerns research traditions that, in different ways, relate to categorization in everyday life, as well as in science. Specifically, the book explores links between two major and intrinsically related scientific disciplines that study these topics. One is cognitive science, in particular, its sub-area that is concerned with human concepts and mental representations of categories. In recent years this discipline has been receiving increasing attention. A crucial issue is how mental representations of categories and concepts are structurally organized. Rosch and Mervis (1975) got things moving: in line with philosophers like Ryle (1951) and Wittgenstein (1953), they questioned the adequacy of the so-called '*classical view*' on concepts that considers concepts as entities with well-defined borderlines and describable by sets of singly necessary and jointly sufficient conditions. Since then, several alternative theoretical views on concepts have been proposed, among them the *prototype view*, the *exemplar view*, the *frame view*, and the '*theory view*'. Proponents of these views have justified them both by means of empirical evidence and by theoretical arguments and counter-examples, some of which, such as Osherson and Smith's (1981) 'pet fish', have become widely known.

This book reviews and analyzes these approaches to concept representation, and also offers new ideas that attempt to overcome their weaknesses. For example, the '*two-tiered view*' tries to reconcile and extend the classical view and other views by recognizing an inherent 'duality' of concept representation: its first component ('tier') captures most stable aspects of a concept, and its second component handles the concept's flexibility and context-dependence through a dynamic inference process.

The other major discipline reflected in the book is classificatory data analysis. Classification problems can be divided into two major types: (1) creating categories based on information about observed objects ('clustering problems'); and (2) inducing general classification rules from descriptions of members and non-members of a given category ('rule induction problems'). Many data-analytic techniques are available to solve these two types of problem.

The field of classificatory data analysis, also called numerical taxonomy, has been active for quite a long time. A milestone in its development was the appearance of the review book by Sokal and Sneath (1963) (see also Sneath and Sokal, 1973). Since then a rich variety of new classification methods has been proposed. Early clustering methods have not been always held in high esteem, as in quite a number of cases they did not rely on a rigorous mathematical model formulation. Since then, however, this situation has evolved, in that various clustering methods were put on a stronger geometric and algebraic (or graph-theoretic) basis. Also, non-parametric, discrete data-analytic methods (that is, methods of 'combinatorial data analysis') were refined, and now enjoy increasing popularity (see, e.g., Arabie and Hubert, 1992; Arabie, Hubert and De Soete, in press; Guénoche and Monjardet, 1987).

The emergence of the field of artificial intelligence has brought a new perspective and produced a largely independent surge of interest in these topics. Studies on categories and concepts have been conducted as components of the now quite broad field of machine learning. Their primary emphasis is on symbolic rather than numeric methods for data characterization, and the methods explored are based on various types of inference: induction, deduction and/or analogy (see, e.g., Kodratoff and Michalski, 1990; Michalski, Carbonell, and Mitchell, 1983, 1986). This research has produced a number of methods and tools for learning general concept descriptions from examples: explanation-based learning (which deductively reformulates the input information), learning and problem solving by analogy, abduction, qualitative discovery, conceptual clustering (which generates classifications with symbolic class descriptions), and other techniques, which go beyond the scope of this book. There is a substantial body of available literature on these efforts (see section 3.2).

This book focuses on bringing together the cognitive research on categories and concepts, and traditional classificatory data analysis. In particular, the book provides an overview of today's most important cognitive-theoretical views on categories and concepts, and a representative sample of methods of classificatory data analysis. An effort is made to test cognitive views against formal models underlying the data-analytic techniques, and vice versa. The next section explores the challenges brought up by this confrontation.

2 Challenges of a confrontation between two research traditions

Bringing together the cognitive and data-analytic research on categorization requires some justification. In general, until now research in each of the fields has evolved largely independently from the other, the only major exception probably being the work of various mathematical psychologists (e.g., Anderson, 1991; Ashby and Gott, 1988; Nosofsky, 1984, 1986; Nosofsky, Clark, and Shin, 1989; Shepard, 1987). On the one hand, cognitive research on categories and concepts generally relies on direct questioning of subjects about the category-related information looked for, rather than on information induced by means of data-analytic techniques. On the other hand, data-analytic techniques have only rarely been investigated with respect to the cognitive validity of their underlying principles. Actually, data analysts seldom make claims about the cognitive validity of their models, let alone the associated algorithms. At best, the outcome of a data-analytic procedure can be considered to correspond to a potential cognitive structure, and whether it corresponds effectively to a cognitive structure can only be proved by means of additional theoretical or empirical evidence.

Moreover, at this time, the connection between cognitive views on categories and classificatory data analysis seems to be less obvious than ever before. For example, probably the most typical procedures of numerical taxonomy are clustering procedures for object by object (dis)similarity data. At the time of the appearance of Sneath and Sokal's book, such procedures looked very promising from a cognitive point of view: similarity seemed to be a straightforward formalization of 'family resemblance', the latter being a central notion in prototype theory, which had recently come into vogue at that time. Yet, in recent cognitive studies the role of similarity as a categorization principle has been questioned (see, e.g., Medin, 1989; Rips, 1989; Rips and Collins, 1992).

Despite all hindrances, we believe that the confrontation of the cognitive and data-analytic research traditions on categorization opens exciting perspectives, and will be very fruitful. A data-analytic questioning of cognitive views can bring cognitive researchers to a more rigorous formalization of their central concepts and classification principles. Furthermore, looking at data-analytic methods from the cognitive theory viewpoint can lead to better insights into the psychological significance of data-analytic model assumptions. Finally, for both disciplines a confrontation could open interesting perspectives for collaborative research. Categorization principles discovered by cognitive theorists could be used to develop new data-analytic procedures; these procedures, in turn, could be used to study, in an inductive and more refined way, categories and concepts.

3 Contents of this book

3.1 What is covered

This book consists of two parts. The first, cognitive part, contains a state-of-the-art overview of current theoretical models of categories and concepts. The basic principles of each model are explained in detail, and are presented along with an overview of supportive theoretical and empirical evidence. Also, for most models the question of which data-analytic methods provide an adequate formalization of the basic model assumptions is explored. The second, data-analytic part contains a representative sample of data-analytic techniques to induce either a set of categories from a data set (possibly along with a corresponding set of category descriptions), or to induce classification rules for one or more categories that are given in advance. For each technique, the formal data-analytic model is spelled out in detail, and references are given to associated algorithms. Most techniques are illustrated by means of applications on hypothetical or real data.

Each part of the book starts with an introductory chapter, which provides a conceptual framework for the domain. In this framework both the models and methods that are covered in the chapters that follow, and relevant models and methods that are not covered, are discussed, together with their interrelations. The two parts are followed by an integrating final chapter. This chapter first examines from a data-analytic viewpoint the consequences of using various cognitive-theoretical models for empirical research. It further investigates the three major aspects of data analysis (namely, data collection, model formulation, and evaluation of results) from a cognitive perspective. Finally, prospects for collaborative research are sketched.

3.2 What is not covered

The focus of the book is on *structural* aspects of concepts and category representations, and traditional methods of data analysis. *Process* aspects are only cursorily dealt with. Therefore, in the cognitive part of the book, human induction and developmental research are not covered. Research results from these fields are only mentioned to the extent that they provide evidence for or against some structural representation. Similarly, the data-analytic part of the book focuses primarily on formal data-analytic models rather than on algorithmic details.

Connectionist or neural network models are also only very cursorily covered in the book (see Barsalou and Hale, Chapter 5; Murphy, Chapter 7). In general, networks operate in a way that is formally similar to that of data-analytic techniques. They can be used to solve most of the inductive problems that are within the scope of classificatory data analysis. Historically, however, network models arose outside the domain of traditional data analysis. The emphasis in their development has usually been on learning algorithms rather than on the mathematical properties of the resulting structural representations. A difference between networks and traditional data analysis is that more claims are made about the cognitive validity of neural networks. In part this is related to the fact that in many network algorithms data are sequentially processed as in human learning processes (see, e.g., Anderson, 1991), whereas in the algorithms of traditional data analysis data are processed all at once ('in batch'). Taken all together, although neural networks could formally be considered to be data-analytic techniques, they differ in several respects from more traditional data analysis. This, together with the fact that the comparison of traditional data analysis and connectionist models has hardly started (Murtagh, in press), caused us not to give an extensive exposition of networks in this book.

As mentioned earlier, the machine-learning methods developed in Artificial Intelligence, though they often address similar problems, represent substantially different approaches and explore a wider range of underlying inference processes, such as induction, deduction, and analogy. It is not possible to cover them together with data-analytic methods in just one book. The exploration of the links and relations between the two groups of methods will require a separate effort. For overviews of relevant work in this area, including research on neural networks and genetic algorithms, the reader may consult, for example, Michalski, Carbonell, and Mitchell (1983, 1986); Laird (1988); Touretzky, Hinton, and Sejnowski (1988); Goldberg (1989); Schafer (1989); Segre (1989); Rivest, Haussler, and Warmuth, 1989; Fulk and Case (1990); Porter and Mooney (1990); Kodratoff and Michalski (1990); Birnbaum and Collins (1991); Michalski and Tecuci (1991, 1992).

3.3 Intended readership and level of the book

The target audience for this book is students of classification and categorization in areas such as cognitive science, data analysis, artificial intelligence, linguistics, medicine, psychiatry, and social sciences. Authors of the individual chapters and the editors have made a special effort to make the book understandable to readers from these disciplines. Chapters have been written in a tutorial fashion. Cognitive science chapters contain explanations of the terminology and basic concepts. The data-analytic chapters emphasize the ideas and the assumptions behind the formal models and methods, rather than algorithmic or technical details. The presupposed level of mathematical foreknowledge is that of an undergraduate course in mathematics/statistics.

References

Anderson, J. R. (1991) 'The adaptive nature of human categorization', *Psychological Review*, 98: 409–29.

Arabie, P. and Hubert, L. J. (1992) 'Combinatorial data analysis', *Annual Review of Psychology*, 43: 169–203.

Arabie, P., Hubert, L. J. and De Soete, G. (eds) (in press) *Clustering and Classification*, Teaneck, NJ: World Scientific.

Ashby, F. G. and Gott, R. E. (1988) 'Decision rules in the perception and categorization of multidimensional stimuli', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14: 33–53.

Birnbaum, L. and Collins, G. (eds) (1991) *Proceedings of the Eighth International Conference on Machine Learning*, Chicago.

Fulk, M. and Case, J. (eds) (1990) *Proceedings of the Third Annual Workshop on Computational Learning Theory*, University of Rochester, NY.

Guénoche, A. and Monjardet, B. (1987) 'Méthodes ordinales et combinatoires en analyse des données', *Mathématiques, Informatique, et Sciences Humaines*, 100: 5–47.

Goldberg, D. E. (1989) *Genetic Algorithms in Search, Optimization, and Machine Learning*, Reading: Addison-Wesley.

Harnad, S. (ed.) (1987) *Categorical Perception*, Cambridge: Cambridge University Press.

Kodratoff, Y. and Michalski, R. S. (eds) (1990) *Machine Learning: An Artificial Intelligence Approach* (vol. III), San Mateo, CA: Morgan Kaufmann.

Laird, J. E. (ed.) (1988) *Proceedings of the Fifth International Conference on Machine Learning*, University of Michigan, Ann Arbor.

Medin, D. L. (1989) 'Concepts and conceptual structure', *American Psychologist*, 44: 1469–81.

Michalski, R. S., Carbonell, J. G., and Mitchell, T. M. (eds) (1983) *Machine Learning: An Artificial Intelligence Approach* (vol. I), San Mateo, CA: Morgan Kaufmann.

Michalski, R. S., Carbonell, J. G., and Mitchell, T. M. (eds) (1986) *Machine Learning: An Artificial Intelligence Approach* (vol. II), San Mateo, CA: Morgan Kaufmann.

Michalski, R. S. and Tecuci, G. (eds) (1991) *Proceedings of the First International Workshop on Multistrategy Learning*, George Mason University, Harper's Ferry, VA.

Michalski, R. S. and Tecuci, G. (eds) (1992) *Machine Learning: A Multistrategy Approach* (vol. IV), San Mateo, CA: Morgan Kaufmann.

Murtagh, F. (in press) 'Neural networks for clustering', in P. Arabie, L. J. Hubert and G. De Soete (eds), *Clustering and Classification*. Teaneck, NJ: World Scientific.

Nosofsky, R. M. (1984) 'Choice, similarity, and the context theory of classification', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10: 104–14.

Nosofsky, R. M. (1986) 'Attention, similarity, and the identification–categorization relationship', *Journal of Experimental Psychology: General*, 115: 39–57.

Nosofsky, R. M., Clark, S. E., and Shin, H. J. (1989) 'Rules and exemplars in categorization, identification, and recognition', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15: 283–304.

Osherson, D. N. and Smith, E. E. (1981) 'On the adequacy of prototype theory as a theory of concepts', *Cognition*, 9: 35–58.

Porter, B. W. and Mooney, R. J. (eds) (1990) *Proceedings of the Seventh International Machine Learning Conference*, Austin, TX.

Rips, L. J. (1989) 'Similarity, typicality, and categorization', in S. Vosniadou and A. Ortony (eds), *Similarity and Analogical Reasoning* (pp. 21–59), Cambridge: Cambridge University Press.

Rips, L. J. and Collins, A. (1992) 'Categories and resemblance', Manuscript submitted for publication.

Rivest, R., Haussler D. and Warmuth, M. (eds) (1989) *Proceedings of the Second Annual Workshop on Computational Learning Theory*, University of Santa Cruz, CA.

Rosch, E. and Mervis, C. (1975) 'Family resemblances: studies in the internal structure of categories', *Cognitive Psychology*, 7: 573–605.

Ryle, G. (1951) 'Polymorphous concepts', *Proceedings of the Aristotelian Society* (Supplementary Series), 25: 65.

Schafer, D. (ed.) (1989) *Proceedings of the Third International Conference on Genetic Algorithms*, George Mason University, Harper's Ferry, VA.

Segre, A. M. (ed.) (1989) *Proceedings of the Sixth International Workshop on Machine Learning*. Cornell University, Ithaca, NY.

Shepard, R. N. (1987) 'Toward a universal law of generalization for psychological science', *Science*, 237: 1317–23.

Sneath, P. H. A. and Sokal, R. R. (1973) *Numerical Taxonomy*, San Francisco: Freeman.

Sokal, R. R. and Sneath, P. H. A. (1963) *Principles of Numerical Taxonomy*, San Francisco: Freeman.

Touretzky, D., Hinton, G., and Sejnowski, T. (eds) (1988) *Proceedings of the 1988 Connectionist Models Summer School*, Carnegie Mellon University, Pittsburgh.

Wittgenstein, L. (1953) *Philosophical investigations*, New York: Macmillan.