Nonmonotonic Reasoning: Logical Foundations of Commonsense

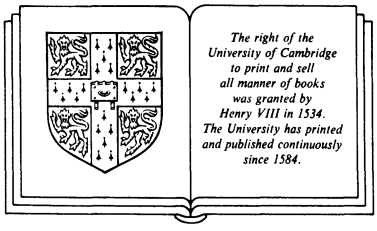
Cambridge Tracts in Theoretical
Computer Science

Titles in the series

1. G. Chaitin *Algorithmic Information Theory*
2. L.C. Paulson *Logic and Computation*
3. M. Spivey *Understanding Z*
4. G. Revesz *Lambda Calculus, Combinators and Logic Programming*
5. A. Ramsay *Formal Methods in Artificial Intelligence*
6. S. Vickers *Topology via Logic*
7. J-Y Girard, Y. Lafont & P. Taylor *Proofs and Types*
8. J. Clifford *Formal Semantics & Pragmatics for Natural Language Processing*
9. M. Winslett *Updating Logical Databases*
10. K. McEvoy & J.V. Tucker (eds) *Theoretical Foundations of VLSI Design*
12. G. Brewka *Nonmonotonic Reasoning*
13. G. Smolka *Logic Programming over Polymorphically Order–Sorted Types*
15. S. Dasgupta *Design Theory and Computer Science*
17. J.C.M. Baeten (ed) *Applications of Process Algebra*
18. J.C.M. Baeten & W. P. Weijland *Process Algebra*

Nonmonotonic Reasoning: Logical Foundations of Commonsense

Gerhard Brewka
GMD,
Sankt Augustin, Germany



CAMBRIDGE UNIVERSITY PRESS
Cambridge
New York Port Chester Melbourne Sydney

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521181303

First published 1991
First paperback edition 2010

A catalogue record for this publication is available from the British Library

ISBN 978-0-521-38394-3 Hardback
ISBN 978-0-521-18130-3 Paperback

CONTENTS

FOREWORD

Systems of whatever kind are becoming increasingly complex. This is certainly true for technical systems. Just think of the unimaginable complexity of the entire 'Space Shuttle' system (including all the computing machinery involved). But it is true of non-technical systems as well, such as social systems (eg. social insurance). Evidence can be encountered every day that this complexity surpasses man's capabilities.

It is unlikely that there is an easy way of solving the resulting problems by attempting to reduce the complexity. So it seems that the only viable alternative consists in providing man with facilities that support his dealing with systems of such complexity. Such an 'interface' system will itself have to become quite complex indeed. Its complexity might be concealed from the user, however, provided its behavior simulates that of a human assistant or expert; for humans have no difficulty in communicating with human 'interfaces'.

In order to approximate the behavior of a human knowledgeable in some special area of expertise, such as providing knowledge about other complex systems, it seems to be necessary to get the system to reason the way human beings do. This is the ultimate goal motivating the material presented in this book.

This goal is far easier stated than achieved. What exactly is the way human beings use reasoning, in the first place? Even if there was some uniform way and we had found it, there would still remain the task of casting it into a formalism suitable for computers. The combination of both these tasks leaves us with what seems the only way of approaching the goal: start with some conjecture about the human way of reasoning; cast it into a formalism; test the formalism's behavior in applications, in comparison with human reasoning; if necessary revise the conjecture and start over again; and so forth.

Research in many places all over the world has already undergone a number of cycles in this procedure already with no end in sight. This situation is well reflected in the present book which introduces as possible candidates, several different formalisms that have been developed by different authors during the last decade. Most of these are extensions of logic, a formalism that evolved in earlier cycles over the past two thous-

and years. One of the book's virtues is its contribution to the clarification of the relationships between those different formalisms. Obviously any future cycle is the more promising the more insight we use in the revision on which it will be based.

The book also offers interesting new ideas towards such a revision. The author's interesting proposal of preferred subtheories is a good example. The book should therefore be valuable for the specialist interested in such new proposals and in the clarified relationships, as well as for the novice in the field who wants to learn about the present state in this area, and hence also for anyone in between these two extremes. I am confident that such readers will find studying this text as rewarding as it has been for me.

W. Bibel,
Lindenau

PREFACE

You can't always get what you want,
but if you try some time you get what you need.
Mick Jagger, Rolling Stones

Yes, you are right. It is probably not the best idea to write a book in a field which is developing as rapidly as nonmonotonic reasoning and where everything is still in a state of flux. Whenever a chapter is finished someone writes another important paper which raises new ideas on that very topic. Not so long ago only a small circle of a few people interested in Artificial Intelligence thought about the problems discussed in this book - and their results were presented at conferences under the heading 'exotic logics'. Now those times are over; there is an explosion of interest. More and more scientists from various areas - philosophers, logicians, linguists, and of course those interested in Artificial Intelligence - detect the importance of nonmonotonic reasoning and contribute their ideas to the field.

So, due to such broad interest, along with the difficulties comes the justification for a book like this. There is no chance to be complete in any sense. I have tried to include in my presentation those approaches which promise to have a bearing on future developments. Of course, there are no clear criteria for selecting the material to be included, and some readers may find their favourite nonmonotonic system is missing. One indicator of quality I used was the appearance of an approach in one of the important Artificial Intelligence journals. And in fact all of the approaches in this book - or at least particular instances of the approaches - have been published in one of them. I hope that this gives my selection at least some objectivity and guarantees that the approaches presented have a certain degree of stability and persistence.

The main goal of this book is to give a broad overview on the state of the art in different fields of research in the area of nonmonotonic reasoning. The book should be understood as a providing a first orientation in a rapidly emerging field. I seek to make it easier to read the original papers, but not superfluous.

After discussing in Chapter 1 various types of nonmonotonic reasoning and their

possible applications I present the major logics for nonmonotonic reasoning:

- modal approaches, in particular McDermott and Doyle's nonmonotonic logic and autoepistemic logic (Chapter 2);
- default logics based on nonmonotonic inference rules (Chapter 3);
- various forms of circumscription (Chapter 4);
- the preferred subtheory approach, a generalization of Poole's logical framework for default reasoning (Chapter 5); and
- Delgrande's conditional logic for default reasoning (Chapter 6).

Theorem proving techniques for these logics are described in Chapter 7. Chapter 8 contains a discussion of a number of formalizations of nonmonotonic inheritance. A pragmatic approach to nonmonotonic reasoning based on nonmonotonic rules is analyzed in Chapter 9. Chapter 10 discusses the achievements in the field in the light of the famous Yale shooting example.

A more detailed overview of the book can be found in Section 1.4.

Two leitmotifs guide the presentation:

- I try to clarify the relation between different research activities in the field which often have been performed independently of each other, and provide some of the missing links; and
- I see many of the different activities as alternative ways of achieving a common goal, the combination of a sound theoretical foundation and efficient computation.

In nonmonotonic reasoning this combination is much harder to achieve than in classical monotonic reasoning. The computational properties of the nonmonotonic logics are awkward, at least if they are general enough to subsume classical first order logic as a special case, because it can be proved that no correct and complete proof procedures exist. But this does not mean that there is no chance at all of doing theoretically sound nonmonotonic reasoning efficiently. In fact it is one of the objectives of this book to present different ways of achieving or, probably better, approximating this goal.

This book does not just give an overview: new results and proposals are presented in some of the chapters. Trying to produce such a combination of a general overview and new material carries some danger: readers interested in an overview may be unhappy about the author's emphasis on his own ideas; readers familiar with the literature and interested in new results may be bored by the many parts they already know. My hope is - on one hand - that the contributions are of enough interest to be contained in an overview, and - on the other hand - that specialists in nonmonotonic

reasoning will not find it too difficult to skip the sections they know all about. The readers of this last type are referred particularly to the following sections.

In Chapter 3 we define a new version of default logic which avoids some counter-intuitive inferences and, most importantly, is cumulative, i.e. the addition of theorems to the premises does not change the derivable formulae.

In Chapter 5 we develop an approach to default reasoning - the preferred subtheory approach - which is based on the notion of preferred maximal consistent subsets of the premises. I define a general framework, show that Poole's system is a particular instance of this framework and describe two generalizations which allow priorities between defaults to be represented. An advantage of this approach is that no non-classical operators or rules have to be introduced. Thus default reasoning can be integrated more easily with other types of commonsense reasoning.

Chapter 7 describes nonmonotonic theorem proving techniques. Our contribution in this chapter is presented in Section 7.4: a modal default proof procedure for a common subset of McDermott and Doyle's nonmonotonic logic and autoepistemic logic. The procedure has been described in earlier papers (Brewka, Wittur 84) (Brewka 86).

Chapter 8 on inheritance systems contains the author's formalization of frame systems, a very popular form of inheritance systems. This formalization is based on circumscription and an earlier version has been published (Brewka 87).

In Chapter 9 we discuss nonmonotonic process systems which are based on the notion of 'current unprovenness' (as distinct from unprovability). I show that the Doyle-style truth maintenance systems commonly used in nonmonotonic process systems can be seen as propositional default logic provers (Brewka 89) (Reinfrank *et al.* 89), which provides this type of truth maintenance systems with a model theoretical semantics. I further present a critique of dependency-directed backtracking, a common consistency maintenance technique, showing that sometimes unwanted results are obtained. Intermediate inconsistencies may lead to the generation of justifications which produce results having nothing to do with the problem solver's knowledge base. Modifications of the algorithms which avoid these problems turn out to be intolerably inefficient.

The book is based on the author's thesis (Brewka 89), but large parts have been rewritten to increase readability and completeness. To keep the length of the text reasonable (and to avoid much repetition) proofs have been omitted but can be found in the original references. Readers should be familiar with classical first order logic. The book uses standard set-theoretical and logical notation with the usual binding rules, i.e. $\forall x.P(x) \wedge Q(x) \vee R(x) \supset S(x)$ reads as $\forall x.([(P(x) \wedge Q(x)) \vee R(x)] \supset S(x))$. The syntax of first order logic and other logics examined is defined in the Appendix.

ACKNOWLEDGEMENTS

Many people have contributed to this book. Only few of them can be named explicitly here. First of all I'd like to thank my thesis supervisors Prof. Dr. B. Neumann, Universität Hamburg, and Prof. Dr. W. Bibel, Technische Hochschule Darmstadt, for many helpful comments and suggestions. GMD, the German National Research Center for Computer Science, provided me with a stimulating environment, a powerful text processing system, and many interesting colleagues. My colleague Ulrich Junker spent many hours discussing all parts of the book with me. I also thank the other members of GMD's Hybrid Inference Systems group, Franco Di Primio, Peter Henne, Gerd Paaß, and Karl Wittur for many interesting discussions. Many members of GMD's Expert Systems group also helped to improve the book, in particular Tom Gordon, Joachim Hertzberg, and Hans Voss.

Thanks also to many participants of the Second International Workshop on Nonmonotonic Reasoning, Grassau, 1988. I benefitted a lot from discussions with Jon Doyle, Oskar Dressler, Vladimir Lifschitz, Paul Morris, David Poole, Michael Reinfrank and many others. The anonymous referees contributed a lot to the final form of the book. I am particularly indebted to David Makinson who always helped to clarify my ideas.

Needless to say that this book would not have been possible without private support of various kinds, most of which was due to three very young ladies named Kristina, Janna and Alena - and their mother.