Statistics for Engineering
and Information Science

Series Editors
M. Jordan, S.L. Lauritzen, J.F. Lawless, V. Nair

Springer Science+Business Media, LLC

Statistics for Engineering and Information Science

Akaike and Kitagawa: The Practice of Time Series Analysis.
Cowell, Dawid, Lauritzen, and Spiegelhalter: Probabilistic Networks and Expert Systems.
Doucet, de Freitas, and Gordon: Sequential Monte Carlo Methods in Practice.
Fine: Feedforward Neural Network Methodology.
Hawkins and Olwell: Cumulative Sum Charts and Charting for Quality Improvement.
Jensen: Bayesian Networks and Decision Graphs.
Marchette: Computer Intrusion Detection and Network Monitoring: A Statistical Viewpoint.
Vapnik: The Nature of Statistical Learning Theory, Second Edition.

Finn V. Jensen

Bayesian Networks and Decision Graphs

With 184 Illustrations

Springer

Finn V. Jensen
Aalborg University
Department of Computer Sciences
Fredrik Bajers Vej 7E
DK-9220 Aalborg Ø
Denmark
fvj@cs.auc.dk

Series Editors
Michael Jordan
Department of Computer Science
University of California, Berkeley
Berkeley, CA 94720
USA

Steffen L. Lauritzen
Department of Mathematical Sciences
Aalborg University
Fredrik Bajers Vej 7G
DK-9220 Aalborg Ø
Denmark

Jerald F. Lawless
Department of Statistics
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada

Vijay Nair
Department of Statistics
University of Michigan
Ann Arbor, MI 48109
USA

Library of Congress Cataloging-in-Publication Data
Jensen, Finn V.
Bayesian networks and decision graphs / Finn V. Jensen.
p. cm. — (Statistics for engineering and information science)
Includes bibliographical references and index.

ISBN 978-1-4757-3504-8 ISBN 978-1-4757-3502-4 (eBook)
DOI 10.1007/978-1-4757-3502-4

1. Bayesian statistical decision theory—Data processing. 2. Machine learning. 3. Neural networks (Computer science). 4. Decision making. I. Title. II. Series.
QA279.5 .J47 2001
519.5′42—dc21 2001020441

Printed on acid-free paper.

9 8 7 6 5 4

www.springer-ny.com

Preface

Ever since the first machines were constructed, artists and scientists have shared a vision of a human-like machine: an autonomous self-moving machine that acts and reasons like a human being. Much effort has been put into this dream, but we are still very far from having androids with even the tiniest similarity to humans.

This does not mean that all of these efforts have been wasted. As a spinoff, we have seen a long series of inventions that can take over very specialized sections of human work. These inventions fall into two categories: machines that can make physical changes in the world and thereby substitute human labor, and machines that can perform activities usually thought of as requiring intellectual skills.

In contemporary science and engineering, we still have this split into two categories. The activity of the first category is mainly concentrated on the construction of robots. The aim is to construct autonomous machines performing "sophisticated" actions such as searching for a cup, finding a way from the office to a lavatory, driving a vehicle in a deserted landscape or walking on two legs. Construction of such robots requires computers to perform certain kinds of artificial intelligence. Basically, it is the kind of intelligence that humans share with most mammals. It involves skills such as visual recognition of items, sound recognition, learning to abstract crucial items from a scene or control of balance and position in 3-D space. Although they are very challenging research tasks, and they certainly require enormous computing power and very sophisticated algorithms, you would not say that these skills are intellectual, and the basis for the activity is the physical appearance of a device that moves. To say it another way: the success criterion is how the algorithms work when controlling a physical machine in real time.

The activity in the second category is basically concerned with reasoning and human activities that we presumably do not share with other animals. The activity is separated from matter. When performed, no changes in the physical world need to take place. The first real success was the automated calculator: a machine that can perform very large and complicated arithmetic calculations. Automated calculation skill is nowadays hardly considered artificial intelligence, and we are now acquainted with computers performing tasks that decades ago were considered highly intellectual (e.g. taking derivatives of functions or reduction of mathematical expressions). When an activity has been understood so well that it can be formalized, it will soon be performed by computers, and gradually we acknowledge that this activity does not really require intelligence.

A branch of research in the second category has to do with reasoning. The first successes were in logical reasoning. Propositional logic is fully formalized, and although some tasks are NP-complete and therefore in some situations intractable for a computer, we have for propositional logic completed the transition from "intellectual task" to "we have computers to do this for us."

Unfortunately, logical reasoning is very limited in scope. It deals with how to infer from propositions that you know are true. Very often you do not know a proposition for certain, but you still need to perform inference from your incomplete and uncertain knowledge. Actually, this is the most common situation for human reasoning. Reasoning under uncertainty is not yet so well understood that it can be formalized entirely for computers. There are several approaches to reasoning under uncertainty. The approach taken in this book is (subjective) probability theory. When the reasoning ends up in a conclusion on a decision, we use *utilities*, and we assume that the decision taken is the one that maximizes the expected utility. In other words, the approach prescribes a certain behavior. We may not always expect this behavior from human beings, and therefore the approach is also termed *normative*. There are alternative approaches to reasoning under uncertainty. Most prominent is *possibility theory*, which in certain contexts is called *fuzzy logic*. The interested reader may consult the wide literature on these approaches.

The aim of normative systems can in short be termed human *wisdom*: **to take decisions on the basis of accumulated and processed experience**. The tasks are of the following types:

- using observations to interpret a situation;
- focusing a search for more information;
- deciding for intervening actions;
- adapting to changing environments;
- learning from experience.

A damping factor for properly exploiting the advances in artificial intelligence has for a long time been the lack of successes in robotics. An autonomous agent that moves, observes, and changes the world must carry a hardly controllable body.

Therefore, the advances have mainly been exploited in *decision support systems*, computer systems which provide advice for humans on highly specialized tasks. With the Internet, the scope of artificial intelligence has widened considerably. The Internet is an ideal nonphysical world for intelligent agents, which are pure spirits without bodies. In the years to come, we will experience a flood of intelligent agents on the Internet, and companies as well as private persons will be able to launch their own agents to explore and collect information on the Internet. Also, we will experience the dark sides of human endeavor. Some agents will destroy, intrude, tell lies and so on, and we will have to defend ourselves against them. Agents will meet agents, and they will have to decide how to treat each other, they will have to learn from past experience, and they will have to adapt to changing environments.

During the 1990s, Bayesian networks and decision graphs attracted a great deal of attention as a framework for building normative systems, not only in research institutions but also in industry. Contrary to most other frameworks for handling uncertainty, a good deal of theoretical insight as well as practical experience is required in order to exploit the opportunities provided by Bayesian networks and decision graphs.

On the other hand, many scientists and engineers wish to exploit the possibilities of normative systems without being experts in the field. This book should meet that demand. It is intended for both classroom use and self-study, and it addresses persons who are interested in exploiting the approach for the construction of decision support systems or bodyless agents.

The theoretical exposition in the book is self-contained, and the mathematical prerequisite is some prior exposure to calculus and elementary graph theory (except for Section 3.4, which requires familiarity with gradients of functions over several variables). The book falls into two parts. In the first part, the emphasis is on gaining practical experience with the use of Baysian networks and decision graphs, and we have assumed that the reader has access to a computer system for handling Bayesian networks and influence diagrams (the exercises marked with an ^{E} require such a system). There are many systems, academic as well as commercial. The following systems can be downloaded for academic use free of charge: Bayesware (www.bayesware.com), BN Toolbox (www.cs.berkeley.edu /~murphyk/Bayes/bnsoft.html), BucketElim (www.ics.uci.edu/~irinar), Genie (www2.sis.pitt.edu/~genie), Hugin (www.hugin.com), Java Bayes (www.cs.cmu.edu/~javabayes/Home), Netica (www.norsys.com), and XBAIES (www.staff.city.ac.uk/~rgc/webpages/xbpage.html). The presentation in this part is based very much on examples, and for overview purposes there is a summary section at the end of each chapter.

The second part is devoted to presenting basic algorithms for normative systems, and the algorithms are exploited to introduce new types of features for decision support systems and bodyless agents. Alhough the exposition is self-contained, it is more mathematically demanding, and it requires that the reader be familiar with working with texts in the mathematical language.

The book is an introduction to Bayesian networks and decision graphs. Many results are not mentioned or just treated superficially. The following textbooks and monographs can be used for further study:

- Judea Pearl, *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann Publishers, 1988.
- Russell Almond, *Graphical Belief Modelling*, Chapman & Hall, 1995.
- Steffen L. Lauritzen, *Graphical Models*, Oxford University Press, 1996.
- Enrique Castillo, José M. Gutiérrez, and Ali S. Hadi, *Expert Systems and Probabilistic Network Models*, Springer-Verlag, 1997.
- Robert G. Cowell, A. Philip Dawid, and Steffen L. Lauritzen, *Probabilistic Networks and Expert Systems*, Springer-Verlag, 1999.

The annual *Conference on Uncertainty in Artificial Intelligence* (www.auai.org) is the main forum for researchers working with Bayesian networks and decision graphs, so the best general references for further reading are the proceedings from these conferences.

Another relevant conference is the biannual *European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty* (ECSQARU). The conference deals with various approaches to uncertainty calculus, and the proceedings are published in the Springer-Verlag series *Lecture Notes in Artificial Intelligence.*

I wish to express my gratitude to several people for ideas during the preparation of the book. First I thank the Ph.D. students at the Research Unit for Decision Support Systems, Olav Bangsø, Søren L. Dittmer, Anders L. Madsen, Thomas D. Nielsen, and Dennis Nilsson, and my colleagues at Aalborg University, Poul S. Eriksen, Uffe Kjærulff, Steffen L. Lauritzen, and Kristian G. Olesen. I also thank the many academic colleagues in the U.S. and Europe with whom I have had the pleasure of exchanging ideas, in particular Linda van der Gaag, Helge Langseth, Claus Skaanning, Marco Valtorta, Jiří Vomlel, Marta Vomlelová, and Yang Xiang. I also thank two years of undergraduate students who had to cope with unfinished drafts of notes for parts of the course on decision support systems. Finally, I am much indebted to Camilla Jørgensen for her very competent LaTeX-coding and for several corrections to my English.

This book is supported by a web site, www.cs.auc.dk/~fvj/bnid.html, which will support readers with solutions and models for selected exercises, a list of errata, special exercises, and other links relevant to the issues in the book.

Aalborg, January 2001 Finn V. Jensen

Contents