

Modeling and Reasoning with Bayesian Networks

This book provides a thorough introduction to the formal foundations and practical applications of Bayesian networks. It provides an extensive discussion of techniques for building Bayesian networks that model real-world situations, including techniques for synthesizing models from design, learning models from data, and debugging models using sensitivity analysis. It also treats exact and approximate inference algorithms at both theoretical and practical levels. The treatment of exact algorithms covers the main inference paradigms based on elimination and conditioning and includes advanced methods for compiling Bayesian networks, time-space tradeoffs, and exploiting local structure of massively connected networks. The treatment of approximate algorithms covers the main inference paradigms based on sampling and optimization and includes influential algorithms such as importance sampling, MCMC, and belief propagation.

The author assumes very little background on the covered subjects, supplying in-depth discussions for theoretically inclined readers and enough practical details to provide an algorithmic cookbook for the system developer.

Adnan Darwiche is a Professor and Chairman of the Computer Science Department at UCLA. He is also the Editor-in-Chief for the *Journal of Artificial Intelligence Research* (JAIR) and a AAAI Fellow.

Cambridge University Press

978-0-521-88438-9 - Modeling and Reasoning with Bayesian Networks

Adnan Darwiche

Frontmatter

[More information](#)

Modeling and Reasoning with Bayesian Networks

Adnan Darwiche

University of California, Los Angeles



CAMBRIDGE
UNIVERSITY PRESS

Cambridge University Press
 978-0-521-88438-9 - Modeling and Reasoning with Bayesian Networks
 Adnan Darwiche
 Frontmatter
[More information](#)

To Jinan, Layla, Sarah, Sikna, and Youssef.

CAMBRIDGE UNIVERSITY PRESS
 Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi

Cambridge University Press
 32 Avenue of the Americas, New York, NY 10013-2473, USA
www.cambridge.org
 Information on this title: www.cambridge.org/9780521884389

© Adnan Darwiche 2009

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2009

Printed in the United States of America

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

Library of Congress Cataloging in Publication data

Darwiche, Adnan, 1966–
 Modeling and reasoning with Bayesian networks / Adnan Darwiche.
 p. cm.

ISBN 978-0-521-88438-9 (hardback)

1. Bayesian statistical decision theory – Graphic methods. 2. Inference. 3. Probabilities. 4. Modeling.

I. Title.

QA279.5.D37 2009

519.5'42–dc22 2008044605

ISBN 978-0-521-88438-9 (hardback)

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party Internet Web sites referred to in this publication and does not guarantee that any content on such Web sites is, or will remain, accurate or appropriate. Information regarding prices, travel timetables, and other factual information given in this work are correct at the time of first printing, but Cambridge University Press does not guarantee the accuracy of such information thereafter.

Contents

Preface	<i>page xi</i>
1 Introduction	1
1.1 Automated Reasoning	1
1.2 Degrees of Belief	4
1.3 Probabilistic Reasoning	6
1.4 Bayesian Networks	8
1.5 What Is Not Covered in This Book	12
2 Propositional Logic	13
2.1 Introduction	13
2.2 Syntax of Propositional Sentences	13
2.3 Semantics of Propositional Sentences	15
2.4 The Monotonicity of Logical Reasoning	18
2.5 Multivalued Variables	19
2.6 Variable Instantiations and Related Notations	20
2.7 Logical Forms	21
Bibliographic Remarks	24
2.8 Exercises	25
3 Probability Calculus	27
3.1 Introduction	27
3.2 Degrees of Belief	27
3.3 Updating Beliefs	30
3.4 Independence	34
3.5 Further Properties of Beliefs	37
3.6 Soft Evidence	39
3.7 Continuous Variables as Soft Evidence	46
Bibliographic Remarks	48
3.8 Exercises	49
4 Bayesian Networks	53
4.1 Introduction	53
4.2 Capturing Independence Graphically	53
4.3 Parameterizing the Independence Structure	56
4.4 Properties of Probabilistic Independence	58
4.5 A Graphical Test of Independence	63
4.6 More on DAGs and Independence	68

vi	CONTENTS	
	Bibliographic Remarks	71
4.7	Exercises	72
4.8	Proofs	75
5	Building Bayesian Networks	76
5.1	Introduction	76
5.2	Reasoning with Bayesian Networks	76
5.3	Modeling with Bayesian Networks	84
5.4	Dealing with Large CPTs	114
5.5	The Significance of Network Parameters	119
	Bibliographic Remarks	121
5.6	Exercises	122
6	Inference by Variable Elimination	126
6.1	Introduction	126
6.2	The Process of Elimination	126
6.3	Factors	128
6.4	Elimination as a Basis for Inference	131
6.5	Computing Prior Marginals	133
6.6	Choosing an Elimination Order	135
6.7	Computing Posterior Marginals	138
6.8	Network Structure and Complexity	141
6.9	Query Structure and Complexity	143
6.10	Bucket Elimination	147
	Bibliographic Remarks	148
6.11	Exercises	148
6.12	Proofs	151
7	Inference by Factor Elimination	152
7.1	Introduction	152
7.2	Factor Elimination	153
7.3	Elimination Trees	155
7.4	Separators and Clusters	157
7.5	A Message-Passing Formulation	159
7.6	The Jointree Connection	164
7.7	The Jointree Algorithm: A Classical View	166
	Bibliographic Remarks	172
7.8	Exercises	173
7.9	Proofs	176
8	Inference by Conditioning	178
8.1	Introduction	178
8.2	Cutset Conditioning	178
8.3	Recursive Conditioning	181
8.4	Any-Space Inference	188
8.5	Decomposition Graphs	189
8.6	The Cache Allocation Problem	192
	Bibliographic Remarks	196

CONTENTS	vii
8.7 Exercises	197
8.8 Proofs	198
9 Models for Graph Decomposition	202
9.1 Introduction	202
9.2 Moral Graphs	202
9.3 Elimination Orders	203
9.4 Jointrees	216
9.5 Dtrees	224
9.6 Triangulated Graphs	229
Bibliographic Remarks	231
9.7 Exercises	232
9.8 Lemmas	234
9.9 Proofs	236
10 Most Likely Instantiations	243
10.1 Introduction	243
10.2 Computing MPE Instantiations	244
10.3 Computing MAP Instantiations	258
Bibliographic Remarks	264
10.4 Exercises	265
10.5 Proofs	267
11 The Complexity of Probabilistic Inference	270
11.1 Introduction	270
11.2 Complexity Classes	271
11.3 Showing Hardness	272
11.4 Showing Membership	274
11.5 Complexity of MAP on Polytrees	275
11.6 Reducing Probability of Evidence to Weighted Model Counting	276
11.7 Reducing MPE to W-MAXSAT	280
Bibliographic Remarks	283
11.8 Exercises	283
11.9 Proofs	284
12 Compiling Bayesian Networks	287
12.1 Introduction	287
12.2 Circuit Semantics	289
12.3 Circuit Propagation	291
12.4 Circuit Compilation	300
Bibliographic Remarks	306
12.5 Exercises	306
12.6 Proofs	309
13 Inference with Local Structure	313
13.1 Introduction	313
13.2 The Impact of Local Structure on Inference Complexity	313
13.3 CNF Encodings with Local Structure	319

viii	CONTENTS	
13.4	Conditioning with Local Structure	323
13.5	Elimination with Local Structure	326
	Bibliographic Remarks	336
13.6	Exercises	337
14	Approximate Inference by Belief Propagation	340
14.1	Introduction	340
14.2	The Belief Propagation Algorithm	340
14.3	Iterative Belief Propagation	343
14.4	The Semantics of IBP	346
14.5	Generalized Belief Propagation	349
14.6	Joingraphs	350
14.7	Iterative Joingraph Propagation	352
14.8	Edge-Deletion Semantics of Belief Propagation	354
	Bibliographic Remarks	364
14.9	Exercises	365
14.10	Proofs	370
15	Approximate Inference by Stochastic Sampling	378
15.1	Introduction	378
15.2	Simulating a Bayesian Network	378
15.3	Expectations	381
15.4	Direct Sampling	385
15.5	Estimating a Conditional Probability	392
15.6	Importance Sampling	393
15.7	Markov Chain Simulation	401
	Bibliographic Remarks	407
15.8	Exercises	408
15.9	Proofs	411
16	Sensitivity Analysis	417
16.1	Introduction	417
16.2	Query Robustness	417
16.3	Query Control	427
	Bibliographic Remarks	433
16.4	Exercises	434
16.5	Proofs	435
17	Learning: The Maximum Likelihood Approach	439
17.1	Introduction	439
17.2	Estimating Parameters from Complete Data	441
17.3	Estimating Parameters from Incomplete Data	444
17.4	Learning Network Structure	455
17.5	Searching for Network Structure	461
	Bibliographic Remarks	466
17.6	Exercises	467
17.7	Proofs	470

CONTENTS		ix
18 Learning: The Bayesian Approach		477
18.1 Introduction		477
18.2 Meta-Networks		479
18.3 Learning with Discrete Parameter Sets		482
18.4 Learning with Continuous Parameter Sets		489
18.5 Learning Network Structure		498
Bibliographic Remarks		504
18.6 Exercises		505
18.7 Proofs		508
A Notation		515
B Concepts from Information Theory		517
C Fixed Point Iterative Methods		520
D Constrained Optimization		523
Bibliography		527
Index		541

Preface

Bayesian networks have received a lot of attention over the last few decades from both scientists and engineers, and across a number of fields, including artificial intelligence (AI), statistics, cognitive science, and philosophy.

Perhaps the largest impact that Bayesian networks have had is on the field of AI, where they were first introduced by Judea Pearl in the midst of a crisis that the field was undergoing in the late 1970s and early 1980s. This crisis was triggered by the surprising realization that a theory of plausible reasoning cannot be based solely on classical logic [McCarthy, 1977], as was strongly believed within the field for at least two decades [McCarthy, 1959]. This discovery has triggered a large number of responses by AI researchers, leading, for example, to the development of a new class of symbolic logics known as non-monotonic logics (e.g., [McCarthy, 1980; Reiter, 1980; McDermott and Doyle, 1980]). Pearl's introduction of Bayesian networks, which is best documented in his book [Pearl, 1988], was actually part of his larger response to these challenges, in which he advocated the use of probability theory as a basis for plausible reasoning and developed Bayesian networks as a practical tool for representing and computing probabilistic beliefs.

From a historical perspective, the earliest traces of using graphical representations of probabilistic information can be found in statistical physics [Gibbs, 1902] and genetics [Wright, 1921]. However, the current formulations of these representations are of a more recent origin and have been contributed by scientists from many fields. In statistics, for example, these representations are studied within the broad class of graphical models, which include Bayesian networks in addition to other representations such as Markov networks and chain graphs [Whittaker, 1990; Edwards, 2000; Lauritzen, 1996; Cowell et al., 1999]. However, the semantics of these models are distinct enough to justify independent treatments. This is why we decided to focus this book on Bayesian networks instead of covering them in the broader context of graphical models, as is done by others [Whittaker, 1990; Edwards, 2000; Lauritzen, 1996; Cowell et al., 1999]. Our coverage is therefore more consistent with the treatments in [Jensen and Nielsen, 2007; Neapolitan, 2004], which are also focused on Bayesian networks.

Even though we approach the subject of Bayesian networks from an AI perspective, we do not delve into the customary philosophical debates that have traditionally surrounded many works on AI. The only exception to this is in the introductory chapter, in which we find it necessary to lay out the subject matter of this book in the context of some historical AI developments. However, in the remaining chapters we proceed with the assumption that the questions being treated are already justified and simply focus on developing the representational and computational techniques needed for addressing them. In doing so, we have taken a great comfort in presenting some of the very classical techniques in ways that may seem unorthodox to the expert. We are driven here by a strong desire to provide the most intuitive explanations, even at the expense of breaking away from norms. We

have also made a special effort to appease the scientist, by our emphasis on justification, and the engineer, through our attention to practical considerations.

There are a number of fashionable and useful topics that we did not cover in this book, which are mentioned in the introductory chapter. Some of these topics were omitted because their in-depth treatment would have significantly increased the length of the book, whereas others were omitted because we believe they conceptually belong somewhere else. In a sense, this book is not meant to be encyclopedic in its coverage of Bayesian networks; rather it is meant to be a focused, thorough treatment of some of the core concepts on modeling and reasoning within this framework.

Acknowledgments

In writing this book, I have benefited a great deal from a large number of individuals who provided help at levels that are too numerous to explicate here. I wish to thank first and foremost members of the automated reasoning group at UCLA for producing quite a bit of the material that is covered in this book, and for their engagement in the writing and proofreading of many of its chapters. In particular, I would like to thank David Allen, Keith Cascio, Hei Chan, Mark Chavira, Arthur Choi, Taylor Curtis, Jinbo Huang, James Park, Knot Pipatsrisawat, and Yuliya Zabiya. Arthur Choi deserves special credit for writing the appendices and most of Chapter 14, for suggesting a number of interesting exercises, and for his dedicated involvement in the last stages of finishing the book. I am also indebted to members of the cognitive systems laboratory at UCLA – Blai Bonet, Ilya Shipster, and Jin Tian – who have thoroughly read and commented on earlier drafts of the book. A number of the students who took the corresponding graduate class at UCLA have also come to the rescue whenever called. I would like to especially thank Alex Dow for writing parts of Chapter 9. Moreover, Jason Aten, Omer Bar-or, Susan Chebotariov, David Chen, Hicham Elmongui, Matt Hayes, Anand Panangadan, Victor Shih, Jae-il Shin, Sam Talaie, and Mike Zaloznyy have all provided detailed feedback on numerous occasions.

I would also like to thank my colleagues who have contributed immensely to this work through either valuable discussions, comments on earlier drafts, or strongly believing in this project and how it was conducted. In this regard, I am indebted to Russ Almond, Bozhena Bidyuk, Hans Bodlaender, Gregory Cooper, Rina Dechter, Marek Druzdzal, David Heckerman, Eric Horvitz, Linda van der Gaag, Hector Geffner, Vibhav Gogate, Russ Greiner, Omid Madani, Ole Mengshoel, Judea Pearl, David Poole, Wojtek Przytula, Silja Renooij, Stuart Russell, Prakash Shenoy, Hector Palacios Verdes, and Changhe Yuan.

Finally, I wish to thank my wife, Jinan, and my daughters, Sarah and Layla, for providing a warm and stimulating environment in which I could conduct my work. This book would not have seen the light without their constant encouragement and support.