Contents

I PROBABILISTIC REASONING **1**

1 Bayesian Reasoning **3**
1.1 Reasoning under uncertainty . . . 3
1.2 Uncertainty in AI . . . 4
1.3 Probability calculus . . . 5
1.3.1 Conditional probability theorems . . . 8
1.3.2 Variables . . . 9
1.4 Interpretations of probability . . . 10
1.5 Bayesian philosophy . . . 11
1.5.1 Bayes' theorem . . . 11
1.5.2 Betting and odds . . . 13
1.5.3 Expected utility . . . 14
1.5.4 Dutch books . . . 15
1.5.5 Bayesian reasoning examples . . . 17
1.6 The goal of Bayesian AI . . . 21
1.7 Achieving Bayesian AI . . . 22
1.8 Are Bayesian networks Bayesian? . . . 22
1.9 Summary . . . 23
1.10 Bibliographic notes . . . 23
1.11 Technical notes . . . 24
1.12 Problems . . . 25

2 Introducing Bayesian Networks **29**
2.1 Introduction . . . 29
2.2 Bayesian network basics . . . 29
2.2.1 Nodes and values . . . 30
2.2.2 Structure . . . 31
2.2.3 Conditional probabilities . . . 32
2.2.4 The Markov property . . . 33
2.3 Reasoning with Bayesian networks . . . 33
2.3.1 Types of reasoning . . . 34
2.3.2 Types of evidence . . . 35
2.3.3 Reasoning with numbers . . . 36
2.4 Understanding Bayesian networks . . . 37
2.4.1 Representing the joint probability distribution . . . 37

2.4.2 Pearl's network construction algorithm 37
2.4.3 Compactness and node ordering 38
2.4.4 Conditional independence 39
2.4.5 d-separation 41
2.5 More examples 43
2.5.1 Earthquake 43
2.5.2 Metastatic cancer 44
2.5.3 Asia 44
2.6 Summary 45
2.7 Bibliographic notes 45
2.8 Problems 47

3 Inference in Bayesian Networks 53
3.1 Introduction 53
3.2 Exact inference in chains 54
3.2.1 Two node network 54
3.2.2 Three node chain 55
3.3 Exact inference in polytrees 56
3.3.1 Kim and Pearl's message passing algorithm 57
3.3.2 Message passing example 60
3.3.3 Algorithm features 62
3.4 Inference with uncertain evidence 62
3.4.1 Using a virtual node 63
3.4.2 Virtual nodes in the message passing algorithm 65
3.5 Exact inference in multiply-connected networks 66
3.5.1 Clustering methods 66
3.5.2 Junction trees 68
3.6 Approximate inference with stochastic simulation 72
3.6.1 Logic sampling 72
3.6.2 Likelihood weighting 74
3.6.3 Markov Chain Monte Carlo (MCMC) 75
3.6.4 Using virtual evidence 75
3.6.5 Assessing approximate inference algorithms 76
3.7 Other computations 77
3.7.1 Belief revision 77
3.7.2 Probability of evidence 78
3.8 Causal inference 79
3.9 Summary 80
3.10 Bibliographic notes 81
3.11 Problems 82

4 Decision networks **89**
4.1 Introduction . . . 89
4.2 Utilities . . . 89
4.3 Decision network basics . . . 91
4.3.1 Node types . . . 91
4.3.2 Football team example . . . 92
4.3.3 Evaluating decision networks . . . 93
4.3.4 Information links . . . 94
4.3.5 Fever example . . . 95
4.3.6 Types of actions . . . 97
4.4 Sequential decision making . . . 98
4.4.1 Test-action combination . . . 98
4.4.2 Real estate investment example . . . 99
4.4.3 Evaluation using a decision tree model . . . 101
4.4.4 Value of information . . . 103
4.4.5 Direct evaluation of decision networks . . . 104
4.5 Dynamic Bayesian networks . . . 104
4.5.1 Nodes, structure and CPTs . . . 105
4.5.2 Reasoning . . . 107
4.5.3 Inference algorithms for DBNs . . . 109
4.6 Dynamic decision networks . . . 110
4.6.1 Mobile robot example . . . 111
4.7 Summary . . . 112
4.8 Bibliographic notes . . . 113
4.9 Problems . . . 114

5 Applications of Bayesian Networks **117**
5.1 Introduction . . . 117
5.2 A Brief Survey of BN Applications . . . 118
5.2.1 Types of reasoning . . . 118
5.2.2 BN structures for medical problems . . . 118
5.2.3 Other medical applications . . . 120
5.2.4 Non-medical applications . . . 120
5.3 Bayesian Poker . . . 122
5.3.1 Five-card stud poker. . . . 123
5.3.2 A decision network for poker . . . 124
5.3.3 Betting with randomization . . . 127
5.3.4 Bluffing . . . 128
5.3.5 Experimental evaluation . . . 129
5.4 Ambulation monitoring and fall detection . . . 129
5.4.1 The domain . . . 129
5.4.2 The DBN model . . . 130
5.4.3 Case-based evaluation . . . 133
5.4.4 An Extended Sensor Model . . . 135
5.5 A Nice Argument Generator (NAG) . . . 136

5.5.1 NAG architecture 136
5.5.2 Example: An asteroid strike 138
5.5.3 The psychology of inference 139
5.5.4 Example: The asteroid strike continues 141
5.5.5 The future of argumentation 141
5.6 Summary 142
5.7 Bibliographic notes 143
5.8 Problems 143

II LEARNING CAUSAL MODELS 147

6 Learning Linear Causal Models 151
6.1 Introduction 151
6.2 Path models 153
6.2.1 Wright's first decomposition rule 155
6.2.2 Parameterizing linear models 159
6.2.3 Learning linear models is complex 159
6.3 Conditional independence learners 161
6.3.1 Markov equivalence 164
6.3.2 PC algorithm 167
6.3.3 Causal discovery versus regression 169
6.4 Summary 170
6.5 Bibliographic notes 170
6.6 Technical notes 170
6.7 Problems 172

7 Learning Probabilities 175
7.1 Introduction 175
7.2 Parameterizing discrete models 176
7.2.1 Parameterizing a binomial model 176
7.2.2 Parameterizing a multinomial model 179
7.3 Incomplete data 181
7.3.1 The Bayesian solution 182
7.3.2 Approximate solutions 182
7.3.3 Incomplete data: summary 187
7.4 Learning local structure 187
7.4.1 Causal interaction 187
7.4.2 Noisy-or connections 188
7.4.3 Classification trees and graphs 189
7.4.4 Logit models 191
7.4.5 Dual model discovery 192
7.5 Summary 192
7.6 Bibliographic notes 193
7.7 Technical notes 193
7.8 Problems 194

8 Learning Discrete Causal Structure **197**
8.1 Introduction . . . 197
8.2 Cooper & Herskovits' K2 . . . 198
8.2.1 Learning variable order . . . 200
8.3 MDL causal discovery . . . 201
8.3.1 Lam and Bacchus's MDL code for causal models . . . 202
8.3.2 Suzuki's MDL code for causal discovery . . . 205
8.4 Metric pattern discovery . . . 205
8.5 CaMML: Causal discovery via MML . . . 206
8.5.1 An MML code for causal structures . . . 207
8.5.2 An MML metric for linear models . . . 209
8.6 CaMML stochastic search . . . 210
8.6.1 Genetic algorithm (GA) search . . . 211
8.6.2 Metropolis search . . . 211
8.6.3 Prior constraints . . . 213
8.6.4 MML models . . . 214
8.6.5 An MML metric for discrete models . . . 215
8.7 Experimental evaluation . . . 215
8.7.1 Qualitative evaluation . . . 216
8.7.2 Quantitative evaluation . . . 216
8.8 Summary . . . 217
8.9 Bibliographic notes . . . 218
8.10 Technical notes . . . 218
8.11 Problems . . . 219

III KNOWLEDGE ENGINEERING **221**

9 Knowledge Engineering with Bayesian Networks **225**
9.1 Introduction . . . 225
9.1.1 Bayesian network modeling tasks . . . 225
9.2 The KEBN process . . . 226
9.2.1 KEBN lifecycle model . . . 226
9.2.2 Prototyping and spiral KEBN . . . 227
9.2.3 Are BNs suitable for the domain problem? . . . 229
9.2.4 Process management . . . 229
9.3 Modeling and elicitation . . . 230
9.3.1 Variables and values . . . 230
9.3.2 Graphical structure . . . 233
9.3.3 Probabilities . . . 241
9.3.4 Local structure . . . 248
9.3.5 Variants of Bayesian networks . . . 251
9.3.6 Modeling example: Missing car. . . . 251
9.3.7 Decision networks . . . 254
9.4 Adaptation . . . 257
9.4.1 Adapting parameters . . . 258

9.4.2 Structural adaptation 260
9.5 Summary 260
9.6 Bibliographic notes 261
9.7 Problems 261

10 Evaluation **263**
10.1 Introduction 263
10.2 Elicitation review 263
10.3 Sensitivity analysis 264
10.3.1 Sensitivity to evidence 264
10.3.2 Sensitivity to changes in parameters 271
10.4 Case-based evaluation 272
10.4.1 Explanation methods 273
10.5 Validation methods 274
10.5.1 Predictive accuracy 275
10.5.2 Expected value 277
10.5.3 Kullback-Leibler divergence 278
10.5.4 Information reward 280
10.5.5 Bayesian information reward 281
10.6 Summary 282
10.7 Bibliographic notes 283
10.8 Technical notes 284
10.9 Problems 286

11 KEBN Case Studies **287**
11.1 Introduction 287
11.2 Bayesian poker revisited 287
11.2.1 The initial prototype 287
11.2.2 Subsequent developments 288
11.2.3 Ongoing Bayesian poker 289
11.2.4 KEBN aspects 290
11.3 An intelligent tutoring system for decimal understanding 290
11.3.1 The ITS domain 291
11.3.2 ITS system architecture 293
11.3.3 Expert elicitation 294
11.3.4 Automated methods 301
11.3.5 Field trial evaluation 303
11.3.6 KEBN aspects 304
11.4 Seabreeze prediction 305
11.4.1 The seabreeze prediction problem 305
11.4.2 The data 306
11.4.3 Bayesian network modeling 307
11.4.4 Experimental evaluation 309
11.4.5 Case study conclusions 311
11.5 Summary 312

A Notation **315**

B Software Packages **317**
B.1 Introduction 317
B.2 History 318
B.3 Murphy's Software Package Survey 318
B.4 BN software 322
B.4.1 Analytica 322
B.4.2 BayesiaLab 323
B.4.3 Bayes Net Toolbox (BNT) 324
B.4.4 GeNIe 325
B.4.5 Hugin 326
B.4.6 JavaBayes 327
B.4.7 MSBNx 328
B.4.8 Netica 328
B.5 Bayesian statistical modeling 329
B.5.1 BUGS 329
B.5.2 First Bayes 330
B.6 Causal discovery programs 330
B.6.1 Bayesware Discoverer 330
B.6.2 CaMML 331
B.6.3 TETRAD 331
B.6.4 WinMine 331