CONTENTS

Preface xv
Acknowledgments xix

1 The challenges of learning 1

1.1 Learning the best path 2
1.2 Areas of application 4
1.3 Major problem classes 12
1.4 The different types of learning 13
1.5 Learning from different communities 16
1.6 Information collection using decision trees 18
1.6.1 A basic decision tree 18
1.6.2 Decision tree for offline learning 20
1.6.3 Decision tree for online learning 21
1.6.4 Discussion 25
1.7 Website and downloadable software 26
1.8 Goals of this book 26
Problems 28

2 Adaptive learning **31**

2.1 The frequentist view 32
2.2 The Bayesian view 33
2.2.1 The updating equations for independent beliefs 34
2.2.2 The expected value of information 36
2.2.3 Updating for correlated normal priors 38
2.2.4 Bayesian updating with an uninformative prior 41
2.3 Updating for non-Gaussian priors 42
2.3.1 The gamma-exponential model 43
2.3.2 The gamma-Poisson model 44
2.3.3 The Pareto-uniform model 45
2.3.4 Models for learning probabilities* 46
2.3.5 Learning an unknown variance* 49
2.4 Monte Carlo simulation 51
2.5 Why does it work?* 54
2.5.1 Derivation of $\tilde{\sigma}$ 54
2.5.2 Derivation of Bayesian updating equations for independent beliefs 55
2.6 Bibliographic notes 57
Problems 57

3 The economics of information **61**

3.1 An elementary information problem 61
3.2 The marginal value of information 65
3.3 An information acquisition problem 68
3.4 Bibliographic notes 70
Problems 70

4 Ranking and selection **71**

4.1 The model 72
4.2 Measurement policies 75
4.2.1 Deterministic vs. sequential policies 75
4.2.2 Optimal sequential policies 76
4.2.3 Heuristic policies 77
4.3 Evaluating policies 81
4.4 More advanced topics* 83
4.4.1 An alternative representation of the probability space 83
4.4.2 Equivalence of using true means and sample estimates 84

4.5 Bibliographic notes 85
Problems 85

5 The knowledge gradient 89

5.1 The knowledge gradient for independent beliefs 90
5.1.1 Computation 91
5.1.2 Some properties of the knowledge gradient 93
5.1.3 The four distributions of learning 94
5.2 The value of information and the S-curve effect 95
5.3 Knowledge gradient for correlated beliefs 98
5.4 The knowledge gradient for some non-Gaussian distributions 103
5.4.1 The gamma-exponential model 104
5.4.2 The gamma-Poisson model 107
5.4.3 The Pareto-uniform model 108
5.4.4 The beta-Bernoulli model 109
5.4.5 Discussion 111
5.5 Relatives of the knowledge gradient 112
5.5.1 Expected improvement 113
5.5.2 Linear loss* 114
5.6 Other issues 116
5.6.1 Anticipatory vs. experiential learning 117
5.6.2 The problem of priors 118
5.6.3 Discussion 120
5.7 Why does it work?* 121
5.7.1 Derivation of the knowledge gradient formula 121
5.8 Bibliographic notes 125
Problems 126

6 Bandit problems 139

6.1 The theory and practice of Gittins indices 141
6.1.1 Gittins indices in the beta-Bernoulli model 142
6.1.2 Gittins indices in the normal-normal model 145
6.1.3 Approximating Gittins indices 147
6.2 Variations of bandit problems 148
6.3 Upper confidence bounding 149
6.4 The knowledge gradient for bandit problems 151
6.4.1 The basic idea 151
6.4.2 Some experimental comparisons 153

6.4.3 Non-normal models 156
6.5 Bibliographic notes 157
Problems 157

7 Elements of a learning problem 163

7.1 The states of our system 164
7.2 Types of decisions 166
7.3 Exogenous information 167
7.4 Transition functions 168
7.5 Objective functions 168
7.5.1 Designing versus controlling 168
7.5.2 Measurement costs 170
7.5.3 Objectives 170
7.6 Evaluating policies 175
7.7 Discussion 177
7.8 Bibliographic notes 178
Problems 178

8 Linear belief models 181

8.1 Applications 182
8.1.1 Maximizing ad clicks 182
8.1.2 Dynamic pricing 184
8.1.3 Housing loans 184
8.1.4 Optimizing dose response 185
8.2 A brief review of linear regression 186
8.2.1 The normal equations 186
8.2.2 Recursive least squares 187
8.2.3 A Bayesian interpretation 188
8.2.4 Generating a prior 189
8.3 The knowledge gradient for a linear model 191
8.4 Application to drug discovery 192
8.5 Application to dynamic pricing 196
8.6 Bibliographic notes 200
Problems 200

9 Subset selection problems 203

9.1 Applications 205
9.2 Choosing a subset using ranking and selection 206

9.2.1 Setting prior means and variances 207
9.2.2 Two strategies for setting prior covariances 208
9.3 Larger sets 209
9.3.1 Using simulation to reduce the problem size 210
9.3.2 Computational issues 212
9.3.3 Experiments 213
9.4 Very large sets 214
9.5 Bibliographic notes 216
Problems 216

10 Optimizing a scalar function 219

10.1 Deterministic measurements 219
10.2 Stochastic measurements 223
10.2.1 The model 223
10.2.2 Finding the posterior distribution 224
10.2.3 Choosing the measurement 226
10.2.4 Discussion 229
10.3 Bibliographic notes 229
Problems 229

11 Optimal bidding 231

11.1 Modeling customer demand 233
11.1.1 Some valuation models 233
11.1.2 The logit model 234
11.2 Bayesian modeling for dynamic pricing 237
11.2.1 A conjugate prior for choosing between two demand curves 237
11.2.2 Moment matching for non-conjugate problems 239
11.2.3 An approximation for the logit model 242
11.3 Bidding strategies 244
11.3.1 An idea from multi-armed bandits 245
11.3.2 Bayes-greedy bidding 245
11.3.3 Numerical illustrations 247
11.4 Why does it work?* 251
11.4.1 Moment matching for Pareto prior 251
11.4.2 Approximating the logistic expectation 252
11.5 Bibliographic notes 253
Problems 254

12 Stopping problems **255**

12.1 Sequential probability ratio test 255
12.2 The secretary problem 260
12.2.1 Setup 261
12.2.2 Solution 263
12.3 Bibliographic notes 266
Problems 266

13 Active learning in statistics **269**

13.1 Deterministic policies 270
13.2 Sequential policies for classification 274
13.2.1 Uncertainty sampling 274
13.2.2 Query by committee 275
13.2.3 Expected error reduction 276
13.3 A variance minimizing policy 277
13.4 Mixtures of Gaussians 279
13.4.1 Estimating parameters 280
13.4.2 Active learning 281
13.5 Bibliographic notes 283

14 Simulation optimization **285**

14.1 Indifference zone selection 287
14.1.1 Batch procedures 288
14.1.2 Sequential procedures 290
14.1.3 The 0-1 procedure: connection to linear loss 291
14.2 Optimal computing budget allocation 292
14.2.1 Indifference-zone version 293
14.2.2 Linear loss version 294
14.2.3 When does it work? 295
14.3 Model-based simulated annealing 296
14.4 Other areas of simulation optimization 298
14.5 Bibliographic notes 299

15 Learning in mathematical programming **301**

15.1 Applications 303
15.1.1 Piloting a hot air balloon 303
15.1.2 Optimizing a portfolio 308

15.1.3 Network problems 309
15.1.4 Discussion 313
15.2 Learning on graphs 313
15.3 Alternative edge selection policies 316
15.4 Learning costs for linear programs* 317
15.5 Bibliographic notes 324

16 Optimizing over continuous measurements 325

16.1 The belief model 327
16.1.1 Updating equations 328
16.1.2 Parameter estimation 330
16.2 Sequential kriging optimization 332
16.3 The knowledge gradient for continuous parameters* 334
16.3.1 Maximizing the knowledge gradient 334
16.3.2 Approximating the knowledge gradient 335
16.3.3 The gradient of the knowledge gradient 336
16.3.4 Maximizing the knowledge gradient 338
16.3.5 The KGCP policy 339
16.4 Efficient global optimization 340
16.5 Experiments 341
16.6 Extension to higher dimensional problems 342
16.7 Bibliographic notes 343

17 Learning with a physical state 345

17.1 Introduction to dynamic programming 347
17.1.1 Approximate dynamic programming 348
17.1.2 The exploration vs. exploitation problem 350
17.1.3 Discussion 351
17.2 Some heuristic learning policies 352
17.3 The local bandit approximation 353
17.4 The knowledge gradient in dynamic programming 355
17.4.1 Generalized learning using basis functions 355
17.4.2 The knowledge gradient 358
17.4.3 Experiments 361
17.5 An expected improvement policy 363
17.6 Bibliographic notes 364

Index 379