1. Record Nr. UNINA9910633910303321

Autore	Dimitrakakis Christos
Titolo	Decision making under uncertainty and reinforcement learning : theory and algorithms / / Christos Dimitrakakis, Ronald Ortner
Pubbl/distr/stampa	Cham, Switzerland : , : Springer, , [2022] ©2022
ISBN	3-031-07614-1
Descrizione fisica	1 online resource (251 pages)
Collana	Intelligent systems reference library ; ; Volume 223
Disciplina	658.403
Soggetti	Decision making - Mathematical models Reinforcement learning Uncertainty
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Intro -- Preface -- Acknowledgements -- Reference -- Contents -- 1 Introduction -- 1.1 Uncertainty and Probability -- 1.2 The Exploration-Exploitation Trade-Off -- 1.3 Decision Theory and Reinforcement Learning -- References -- 2 Subjective Probability and Utility -- 2.1 Subjective Probability -- 2.1.1 Relative Likelihood -- 2.1.2 Subjective Probability Assumptions -- 2.1.3 Assigning Unique Probabilities* -- 2.1.4 Conditional Likelihoods -- 2.1.5 Probability Elicitation -- 2.2 Updating Beliefs: Bayes' Theorem -- 2.3 Utility Theory -- 2.3.1 Rewards and Preferences -- 2.3.2 Preferences Among Distributions -- 2.3.3 Utility -- 2.3.4 Measuring Utility* -- 2.3.5 Convex and Concave Utility Functions -- 2.4 Exercises -- Reference -- 3 Decision Problems -- 3.1 Introduction -- 3.2 Rewards that Depend on the Outcome of an Experiment -- 3.2.1 Formalisation of the Problem Setting -- 3.2.2 Decision Diagrams -- 3.2.3 Statistical Estimation* -- 3.3 Bayes Decisions -- 3.3.1 Convexity of the Bayes-Optimal Utility* -- 3.4 Statistical and Strategic Decision Making -- 3.4.1 Alternative Notions of Optimality -- 3.4.2 Solving Minimax Problems* -- 3.4.3 Two-Player Games -- 3.5 Decision Problems with Observations -- 3.5.1 Maximizing Utility When Making Observations -- 3.5.2 Bayes Decision Rules -- 3.5.3 Decision Problems in Classification -- 3.5.4 Calculating Posteriors -- 3.6 Summary -- 3.7 Exercises -- 3.7.1 Problems with No

Observations -- 3.7.2 Problems with Observations -- 3.7.3 An Insurance Problem -- 3.7.4 Medical Diagnosis -- References -- 4 Estimation -- 4.1 Introduction -- 4.2 Sufficient Statistics -- 4.2.1 Sufficient Statistics -- 4.2.2 Exponential Families -- 4.3 Conjugate Priors -- 4.3.1 Bernoulli-Beta Conjugate Pair -- 4.3.2 Conjugates for the Normal Distribution -- 4.3.3 Conjugates for Multivariate Distributions -- 4.4 Credible Intervals.
4.5 Concentration Inequalities -- 4.5.1 Chernoff-Hoeffding Bounds -- 4.6 Approximate Bayesian Approaches -- 4.6.1 Monte Carlo Inference -- 4.6.2 Approximate Bayesian Computation -- 4.6.3 Analytic Approximations of the Posterior -- 4.6.4 Maximum Likelihood and Empirical Bayes Methods -- References -- 5 Sequential Sampling -- 5.1 Gains From Sequential Sampling -- 5.1.1 An Example: Sampling with Costs -- 5.2 Optimal Sequential Sampling Procedures -- 5.2.1 Multi-stage Problems -- 5.2.2 Backwards Induction for Bounded Procedures -- 5.2.3 Unbounded Sequential Decision Procedures -- 5.2.4 The Sequential Probability Ratio Test -- 5.2.5 Wald's Theorem -- 5.3 Martingales -- 5.4 Markov Processes -- 5.5 Exercises -- 6 Experiment Design and Markov Decision Processes -- 6.1 Introduction -- 6.2 Bandit Problems -- 6.2.1 An Example: Bernoulli Bandits -- 6.2.2 Decision-Theoretic Bandit Process -- 6.3 Markov Decision Processes and Reinforcement Learning -- 6.3.1 Value Functions -- 6.4 Finite Horizon, Undiscounted Problems -- 6.4.1 Direct Policy Evaluation -- 6.4.2 Backwards Induction Policy Evaluation -- 6.4.3 Backwards Induction Policy Optimization -- 6.5 Infinite-Horizon -- 6.5.1 Examples -- 6.5.2 Markov Chain Theory for Discounted Problems -- 6.5.3 Optimality Equations -- 6.5.4 MDP Algorithms for Infinite Horizon and Discounted Rewards -- 6.6 Optimality Criteria -- 6.7 Summary -- 6.8 Further Reading -- 6.9 Exercises -- 6.9.1 MDP Theory -- 6.9.2 Automatic Algorithm Selection -- 6.9.3 Scheduling -- 6.9.4 General Questions -- References -- 7 Simulation-Based Algorithms -- 7.1 Introduction -- 7.1.1 The Robbins-Monro Approximation -- 7.1.2 The Theory of the Approximation -- 7.2 Dynamic Problems -- 7.2.1 Monte Carlo Policy Evaluation and Iteration -- 7.2.2 Monte Carlo Updates -- 7.2.3 Temporal Difference Methods -- 7.2.4 Stochastic Value Iteration Methods.
7.3 Discussion -- 7.4 Exercises -- References -- 8 Approximate Representations -- 8.1 Introduction -- 8.1.1 Fitting a Value Function -- 8.1.2 Fitting a Policy -- 8.1.3 Features -- 8.1.4 Estimation Building Blocks -- 8.1.5 The Value Estimation Step -- 8.1.6 Policy Estimation -- 8.2 Approximate Policy Iteration (API) -- 8.2.1 Error Bounds for Approximate Value Functions -- 8.2.2 Rollout-Based Policy Iteration Methods -- 8.2.3 Least Squares Methods -- 8.3 Approximate Value Iteration -- 8.3.1 Approximate Backwards Induction -- 8.3.2 State Aggregation -- 8.3.3 Representative State Approximation -- 8.3.4 Bellman Error Methods -- 8.4 Policy Gradient -- 8.4.1 Stochastic Policy Gradient -- 8.4.2 Practical Considerations -- 8.5 Examples -- 8.6 Further Reading -- 8.7 Exercises -- References -- 9 Bayesian Reinforcement Learning -- 9.1 Introduction -- 9.1.1 Acting in Unknown MDPs -- 9.1.2 Updating the Belief -- 9.2 Finding Bayes-Optimal Policies -- 9.2.1 The Expected MDP Heuristic -- 9.2.2 The Maximum MDP Heuristic -- 9.2.3 Bayesian Policy Gradient -- 9.2.4 The Belief-Augmented MDP -- 9.2.5 Branch and Bound -- 9.2.6 Bounds on the Expected Utility -- 9.2.7 Estimating Lower Bounds on the Value Function with Backwards Induction -- 9.2.8 Further Reading -- 9.3 Bayesian Methods in Continuous Spaces -- 9.3.1 Linear-Gaussian Transition Models -- 9.3.2 Approximate Dynamic Programming -- 9.4 Partially Observable Markov Decision Processes -- 9.4.1 Solving Known

POMDPs -- 9.4.2 Solving Unknown POMDPs -- 9.5 Relations Between Different Settings -- 9.6 Exercises -- References -- 10 Distribution-Free Reinforcement Learning -- 10.1 Introduction -- 10.2 Finite Stochastic Bandit Problems -- 10.2.1 The UCB1 Algorithm -- 10.2.2 Non i.i.d. Rewards -- 10.3 Reinforcement Learning in MDPs -- 10.3.1 An Upper-Confidence Bound Algorithm -- 10.3.2 Bibliographical Remarks -- References.
11 Conclusion -- Appendix Symbols -- Appendix Index -- Index.