

1. Record Nr.	UNINA9910829915903321
Titolo	Reinforcement learning and approximate dynamic programming for feedback control // edited by Frank L. Lewis, Derong Liu
Pubbl/distr/stampa	Hoboken, New Jersey : , : John Wiley & Sons, Inc., , [2013] [Piscataway, New Jersey] : , : IEEE Xplore, , [2013]
ISBN	1-299-18985-7 1-118-45393-X
Descrizione fisica	1 online resource (643 p.)
Collana	IEEE Press series on computational intelligence ; ; 17
Classificazione	TEC008000
Disciplina	003.5 003/.5 006.3
Soggetti	Reinforcement learning Feedback control systems
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Description based upon print version of record.
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	PREFACE xix -- CONTRIBUTORS xxiii -- PART I FEEDBACK CONTROL USING RL AND ADP -- 1. Reinforcement Learning and Approximate Dynamic Programming (RLADP)-Foundations, Common Misconceptions, and the Challenges Ahead 3 -- Paul J. Werbos -- 1.1 Introduction 3 -- 1.2 What is RLADP? 4 -- 1.3 Some Basic Challenges in Implementing ADP 14 -- 2. Stable Adaptive Neural Control of Partially Observable Dynamic Systems 31 -- J. Nate Knight and Charles W. Anderson -- 2.1 Introduction 31 -- 2.2 Background 32 -- 2.3 Stability Bias 35 -- 2.4 Example Application 38 -- 3. Optimal Control of Unknown Nonlinear Discrete-Time Systems Using the Iterative Globalized Dual Heuristic Programming Algorithm 52 -- Derong Liu and Ding Wang -- 3.1 Background Material 53 -- 3.2 Neuro-Optimal Control Scheme Based on the Iterative ADP Algorithm 55 -- 3.3 Generalization 67 -- 3.4 Simulation Studies 68 -- 3.5 Summary 74 -- 4. Learning and Optimization in Hierarchical Adaptive Critic Design 78 -- Haibo He, Zhen Ni, and Dongbin Zhao -- 4.1 Introduction 78 -- 4.2 Hierarchical ADP Architecture with Multiple-Goal Representation 80 -- 4.3 Case Study: The Ball-and-Beam System 87 -- 4.4 Conclusions and Future

Work 94 -- 5. Single Network Adaptive Critics Networks-Development, Analysis, and Applications 98 -- Jie Ding, Ali Heydari, and S.N. Balakrishnan -- 5.1 Introduction 98 -- 5.2 Approximate Dynamic Programming 100 -- 5.3 SNAC 102 -- 5.4 J-SNAC 104 -- 5.5 Finite-SNAC 108 -- 5.6 Conclusions 116 -- 6. Linearly Solvable Optimal Control 119 -- K. Dvijotham and E. Todorov -- 6.1 Introduction 119 -- 6.2 Linearly Solvable Optimal Control Problems 123 -- 6.3 Extension to Risk-Sensitive Control and Game Theory 130 -- 6.4 Properties and Algorithms 134 -- 6.5 Conclusions and Future Work 139 -- 7. Approximating Optimal Control with Value Gradient Learning 142 -- Michael Fairbank, Danil Prokhorov, and Eduardo Alonso -- 7.1 Introduction 142 -- 7.2 Value Gradient Learning and BPTT Algorithms 144 -- 7.3 A Convergence Proof for VGL(1) for Control with Function Approximation 148. 7.4 Vertical Lander Experiment 154 -- 7.5 Conclusions 159 -- 8. A Constrained Backpropagation Approach to Function Approximation and Approximate Dynamic Programming 162 -- Silvia Ferrari, Keith Rudd, and Gianluca Di Muro -- 8.1 Background 163 -- 8.2 Constrained Backpropagation (CPROP) Approach 163 -- 8.3 Solution of Partial Differential Equations in Nonstationary Environments 170 -- 8.4 Preserving Prior Knowledge in Exploratory Adaptive Critic Designs 174 -- 8.5 Summary 179 -- 9. Toward Design of Nonlinear ADP Learning Controllers with Performance Assurance 182 -- Jennie Si, Lei Yang, Chao Lu, Kostas S. Tsakalis, and Armando A. Rodriguez -- 9.1 Introduction 183 -- 9.2 Direct Heuristic Dynamic Programming 184 -- 9.3 A Control Theoretic View on the Direct HDP 186 -- 9.4 Direct HDP Design with Improved Performance Case 1-Design Guided by a Priori LQR Information 193 -- 9.5 Direct HDP Design with Improved Performance Case 2-Direct HDP for Coordinated Damping Control of Low-Frequency Oscillation 198 -- 9.6 Summary 201 -- 10. Reinforcement Learning Control with Time-Dependent Agent Dynamics 203 -- Kenton Kirkpatrick and John Valasek -- 10.1 Introduction 203 -- 10.2 Q-Learning 205 -- 10.3 Sampled Data Q-Learning 209 -- 10.4 System Dynamics Approximation 213 -- 10.5 Closing Remarks 218 -- 11. Online Optimal Control of Nonaffine Nonlinear Discrete-Time Systems without Using Value and Policy Iterations 221 -- Hassan Zargarzadeh, Qinmin Yang, and S. Jagannathan -- 11.1 Introduction 221 -- 11.2 Background 224 -- 11.3 Reinforcement Learning Based Control 225 -- 11.4 Time-Based Adaptive Dynamic Programming-Based Optimal Control 234 -- 11.5 Simulation Result 247 -- 12. An Actor-Critic-Identifier Architecture for Adaptive Approximate Optimal Control 258 -- S. Bhasin, R. Kamalapurkar, M. Johnson, K.G. Vamvoudakis, F.L. Lewis, and W.E. Dixon -- 12.1 Introduction 259 -- 12.2 Actor-Critic-Identifier Architecture for HJB Approximation 260 -- 12.3 Actor-Critic Design 263 -- 12.4 Identifier Design 264. 12.5 Convergence and Stability Analysis 270 -- 12.6 Simulation 274 -- 12.7 Conclusion 275 -- 13. Robust Adaptive Dynamic Programming 281 -- Yu Jiang and Zhong-Ping Jiang -- 13.1 Introduction 281 -- 13.2 Optimality Versus Robustness 283 -- 13.3 Robust-ADP Design for Disturbance Attenuation 288 -- 13.4 Robust-ADP for Partial-State Feedback Control 292 -- 13.5 Applications 296 -- 13.6 Summary 300 -- PART II LEARNING AND CONTROL IN MULTIAGENT GAMES -- 14. Hybrid Learning in Stochastic Games and Its Application in Network Security 305 -- Quanyan Zhu, Hamidou Tembine, and Tamer Basar -- 14.1 Introduction 305 -- 14.2 Two-Person Game 308 -- 14.3 Learning in NZSGs 310 -- 14.4 Main Results 314 -- 14.5 Security Application 322 -- 14.6 Conclusions and Future Works 326 -- 15. Integral Reinforcement Learning for Online Computation of Nash Strategies of

Nonzero-Sum Differential Games 330 -- Draguna Vrabie and F.L. Lewis -- 15.1 Introduction 331 -- 15.2 Two-Player Games and Integral Reinforcement Learning 333 -- 15.3 Continuous-Time Value Iteration to Solve the Riccati Equation 337 -- 15.4 Online Algorithm to Solve Nonzero-Sum Games 339 -- 15.5 Analysis of the Online Learning Algorithm for NZS Games 342 -- 15.6 Simulation Result for the Online Game Algorithm 345 -- 15.7 Conclusion 347 -- 16. Online Learning Algorithms for Optimal Control and Dynamic Games 350 -- Kyriakos G. Vamvoudakis and Frank L. Lewis -- 16.1 Introduction 350 -- 16.2 Optimal Control and the Continuous Time Hamilton-Jacobi-Bellman Equation 352 -- 16.3 Online Solution of Nonlinear Two-Player Zero-Sum Games and Hamilton-Jacobi-Isaacs Equation 360 -- 16.4 Online Solution of Nonlinear Nonzero-Sum Games and Coupled Hamilton-Jacobi Equations 366 -- PART III FOUNDATIONS IN MDP AND RL -- 17. Lambda-Policy Iteration: A Review and a New Implementation 381 -- Dimitri P. Bertsekas -- 17.1 Introduction 381 -- 17.2 Lambda-Policy Iteration without Cost Function Approximation 386 -- -- 17.3 Approximate Policy Evaluation Using Projected Equations 388. 17.4 Lambda-Policy Iteration with Cost Function Approximation 395 -- 17.5 Conclusions 406 -- 18. Optimal Learning and Approximate Dynamic Programming 410 -- Warren B. Powell and Ilya O. Ryzhov -- 18.1 Introduction 410 -- 18.2 Modeling 411 -- 18.3 The Four Classes of Policies 412 -- 18.4 Basic Learning Policies for Policy Search 416 -- 18.5 Optimal Learning Policies for Policy Search 421 -- 18.6 Learning with a Physical State 427 -- 19. An Introduction to Event-Based Optimization: Theory and Applications 432 -- Xi-Ren Cao, Yanjia Zhao, Qing-Shan Jia, and Qianchuan Zhao -- 19.1 Introduction 432 -- 19.2 Literature Review 433 -- 19.3 Problem Formulation 434 -- 19.4 Policy Iteration for EBO 435 -- 19.5 Example: Material Handling Problem 441 -- 19.6 Conclusions 448 -- 20. Bounds for Markov Decision Processes 452 -- Vijay V. Desai, Vivek F. Farias, and Ciamac C. Moallemi -- 20.1 Introduction 452 -- 20.2 Problem Formulation 455 -- 20.3 The Linear Programming Approach 456 -- 20.4 The Martingale Duality Approach 458 -- 20.5 The Pathwise Optimization Method 461 -- 20.6 Applications 463 -- 20.7 Conclusion 470 -- 21. Approximate Dynamic Programming and Backpropagation on Timescales 474 -- John Seiffertt and Donald Wunsch -- 21.1 Introduction: Timescales Fundamentals 474 -- 21.2 Dynamic Programming 479 -- 21.3 Backpropagation 485 -- 21.4 Conclusions 492 -- 22. A Survey of Optimistic Planning in Markov Decision Processes 494 -- Lucian Busoniu, Remi Munos, and Robert Babuska -- 22.1 Introduction 494 -- 22.2 Optimistic Online Optimization 497 -- 22.3 Optimistic Planning Algorithms 500 -- 22.4 Related Planning Algorithms 509 -- 22.5 Numerical Example 510 -- 23. Adaptive Feature Pursuit: Online Adaptation of Features in Reinforcement Learning 517 -- Shalabh Bhatnagar, Vivek S. Borkar, and L.A. Prashanth -- 23.1 Introduction 517 -- 23.2 The Framework 520 -- 23.3 The Feature Adaptation Scheme 522 -- 23.4 Convergence Analysis 525 -- 23.5 Application to Traffic Signal Control 527. 23.6 Conclusions 532 -- 24. Feature Selection for Neuro-Dynamic Programming 535 -- Dayu Huang, W. Chen, P. Mehta, S. Meyn, and A. Surana -- 24.1 Introduction 535 -- 24.2 Optimality Equations 536 -- 24.3 Neuro-Dynamic Algorithms 542 -- 24.4 Fluid Models 551 -- 24.5 Diffusion Models 554 -- 24.6 Mean Field Games 556 -- 24.7 Conclusions 557 -- 25. Approximate Dynamic Programming for Optimizing Oil Production 560 -- Zheng Wen, Louis J. Durlofsky, Benjamin Van Roy, and Khalid Aziz -- 25.1 Introduction 560 -- 25.2 Petroleum Reservoir Production Optimization Problem 562 -- 25.3 Review of Dynamic Programming and Approximate Dynamic

Programming 564 -- 25.4 Approximate Dynamic Programming Algorithm for Reservoir Production Optimization 566 -- 25.5 Simulation Results 573 -- 25.6 Concluding Remarks 578 -- 23.6 Conclusions 532 -- 24. Feature Selection for Neuro-Dynamic Programming 535 -- Dayu Huang, W. Chen, P. Mehta, S. Meyn, and A. Surana -- 24.1 Introduction 535 -- 24.2 Optimality Equations 536 -- 24.3 Neuro-Dynamic Algorithms 542 -- 24.4 Fluid Models 551 -- 24.5 Diffusion Models 554 -- 24.6 Mean Field Games 556 -- 24.7 Conclusions 557 -- 25. Approximate Dynamic Programming for Optimizing Oil Production 560 -- Zheng Wen, Louis J. Durllofsky, Benjamin Van Roy, and Khalid Aziz -- 25.1 Introduction 560 -- 25.2 Petroleum Reservoir Production Optimization Problem 562 -- 25.3 Review of Dynamic Programming and Approximate Dynamic Programming 564 -- 25.4 Approximate Dynamic Programming Algorithm for Reservoir Production Optimization 566 -- 25.5 Simulation Results 573 -- 25.6 Concluding Remarks 578 -- 26. A Learning Strategy for Source Tracking in Unstructured Environments 582 -- Titus Appel, Rafael Fierro, Brandon Rohrer, Ron Lumia, and John Wood -- 26.1 Introduction 582 -- 26.2 Reinforcement Learning 583 -- 26.3 Light-Following Robot 589 -- 26.4 Simulation Results 592 -- 26.5 Experimental Results 595 -- 26.6 Conclusions and Future Work 599 -- References 599 -- INDEX 601.

Sommario/riassunto

"Reinforcement learning (RL) and adaptive dynamic programming (ADP) has been one of the most critical research fields in science and engineering for modern complex systems. This book describes the latest RL and ADP techniques for decision and control in human engineered systems, covering both single player decision and control and multi-player games. Edited by the pioneers of RL and ADP research, the book brings together ideas and methods from many fields and provides an important and timely guidance on controlling a wide variety of systems, such as robots, industrial processes, and economic decision-making"--

"Reinforcement learning and adaptive control can be useful for controlling a wide variety of systems including robots, industrial processes, and economical decision making"--
