

1. Record Nr.	UNINA9910488722303321
Titolo	Handbook of reinforcement learning and control // Kyriakos G. Vamvoudakis [and three others], editors
Pubbl/distr/stampa	Cham, Switzerland : , : Springer, , [2021] ©2021
ISBN	3-030-60990-1
Descrizione fisica	1 online resource (839 pages)
Collana	Studies in Systems, Decision and Control ; ; Volume 325
Disciplina	006.31
Soggetti	Reinforcement learning Automatic control - Sensitivity Aprenentatge per reforç (Intel·ligència artificial) Control automàtic Llibres electrònics
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- Preface -- Contents -- Part I Theory of Reinforcement Learning for Model-Free and Model-Based Control and Games -- 1 What May Lie Ahead in Reinforcement Learning -- References -- 2 Reinforcement Learning for Distributed Control and Multi-player Games -- 2.1 Introduction -- 2.2 Optimal Control of Continuous-Time Systems -- 2.2.1 IRL with Experience Replay Learning Technique ch2Modares2014Automatica,ch2Kamalapurkar2016 -- 2.2.2 mathcalHinfity Control of CT Systems -- 2.3 Nash Games -- 2.4 Graphical Games -- 2.4.1 Off-Policy RL for Graphical Games -- 2.5 Output Synchronization of Multi-agent Systems -- 2.6 Conclusion and Open Research Directions -- References -- 3 From Reinforcement Learning to Optimal Control: A Unified Framework for Sequential Decisions -- 3.1 Introduction -- 3.2 The Communities of Sequential Decisions -- 3.3 Stochastic Optimal Control Versus Reinforcement Learning -- 3.3.1 Stochastic Control -- 3.3.2 Reinforcement Learning -- 3.3.3 A Critique of the MDP Modeling Framework -- 3.3.4 Bridging Optimal Control and Reinforcement Learning -- 3.4 The Universal Modeling Framework -- 3.4.1 Dimensions of a Sequential Decision

Model -- 3.4.2 State Variables -- 3.4.3 Objective Functions -- 3.4.4
Notes -- 3.5 Energy Storage Illustration -- 3.5.1 A Basic Energy Storage
Problem -- 3.5.2 With a Time-Series Price Model -- 3.5.3 With Passive
Learning -- 3.5.4 With Active Learning -- 3.5.5 With Rolling Forecasts
-- 3.5.6 Remarks -- 3.6 Designing Policies -- 3.6.1 Policy Search --
3.6.2 Lookahead Approximations -- 3.6.3 Hybrid Policies -- 3.6.4
Remarks -- 3.6.5 Stochastic Control, Reinforcement Learning, and the
Four Classes of Policies -- 3.7 Policies for Energy Storage -- 3.8
Extension to Multi-agent Systems -- 3.9 Observations -- References --
4 Fundamental Design Principles for Reinforcement Learning
Algorithms -- 4.1 Introduction.
4.1.1 Stochastic Approximation and Reinforcement Learning -- 4.1.2
Sample Complexity Bounds -- 4.1.3 What Will You Find in This
Chapter? -- 4.1.4 Literature Survey -- 4.2 Stochastic Approximation:
New and Old Tricks -- 4.2.1 What is Stochastic Approximation? --
4.2.2 Stochastic Approximation and Learning -- 4.2.3 Stability and
Convergence -- 4.2.4 Zap-Stochastic Approximation -- 4.2.5 Rates of
Convergence -- 4.2.6 Optimal Convergence Rate -- 4.2.7 TD and LSTD
Algorithms -- 4.3 Zap Q-Learning: Fastest Convergent Q-Learning --
4.3.1 Markov Decision Processes -- 4.3.2 Value Functions and the
Bellman Equation -- 4.3.3 Q-Learning -- 4.3.4 Tabular Q-Learning --
4.3.5 Convergence and Rate of Convergence -- 4.3.6 Zap Q-Learning
-- 4.4 Numerical Results -- 4.4.1 Finite State-Action MDP -- 4.4.2
Optimal Stopping in Finance -- 4.5 Zap-Q with Nonlinear Function
Approximation -- 4.5.1 Choosing the Eligibility Vectors -- 4.5.2
Theory and Challenges -- 4.5.3 Regularized Zap-Q -- 4.6 Conclusions
and Future Work -- References -- 5 Mixed Density Methods for
Approximate Dynamic Programming -- 5.1 Introduction -- 5.2
Unconstrained Affine-Quadratic Regulator -- 5.3 Regional Model-
Based Reinforcement Learning -- 5.3.1 Preliminaries -- 5.3.2 Regional
Value Function Approximation -- 5.3.3 Bellman Error -- 5.3.4 Actor
and Critic Update Laws -- 5.3.5 Stability Analysis -- 5.3.6 Summary --
5.4 Local (State-Following) Model-Based Reinforcement Learning --
5.4.1 StaF Kernel Functions -- 5.4.2 Local Value Function
Approximation -- 5.4.3 Actor and Critic Update Laws -- 5.4.4 Analysis
-- 5.4.5 Stability Analysis -- 5.4.6 Summary -- 5.5 Combining
Regional and Local State-Following Approximations -- 5.6
Reinforcement Learning with Sparse Bellman Error Extrapolation -- 5.7
Conclusion -- References -- 6 Model-Free Linear Quadratic Regulator.
6.1 Introduction to a Model-Free LQR Problem -- 6.2 A Gradient-Based
Random Search Method -- 6.3 Main Results -- 6.4 Proof Sketch --
6.4.1 Controlling the Bias -- 6.4.2 Correlation of $f(K)$ and $f(K)$ --
6.5 An Example -- 6.6 Thoughts and Outlook -- References -- Part
II Constraint-Driven and Verified RL -- 7 Adaptive Dynamic
Programming in the Hamiltonian-Driven Framework -- 7.1
Introduction -- 7.1.1 Literature Review -- 7.1.2 Motivation -- 7.1.3
Structure -- 7.2 Problem Statement -- 7.3 Hamiltonian-Driven
Framework -- 7.3.1 Policy Evaluation -- 7.3.2 Policy Comparison --
7.3.3 Policy Improvement -- 7.4 Discussions on the Hamiltonian-
Driven ADP -- 7.4.1 Implementation with Critic-Only Structure -- 7.4.2
Connection to Temporal Difference Learning -- 7.4.3 Connection to
Value Gradient Learning -- 7.5 Simulation Study -- 7.6 Conclusion --
References -- 8 Reinforcement Learning for Optimal Adaptive Control
of Time Delay Systems -- 8.1 Introduction -- 8.2 Problem Description
-- 8.3 Extended State Augmentation -- 8.4 State Feedback Q-Learning
Control of Time Delay Systems -- 8.5 Output Feedback Q-Learning
Control of Time Delay Systems -- 8.6 Simulation Results -- 8.7
Conclusions -- References -- 9 Optimal Adaptive Control of Partially

Uncertain Linear Continuous-Time Systems with State Delay -- 9.1 Introduction -- 9.2 Problem Statement -- 9.3 Linear Quadratic Regulator Design -- 9.3.1 Periodic Sampled Feedback -- 9.3.2 Event Sampled Feedback -- 9.4 Optimal Adaptive Control -- 9.4.1 Periodic Sampled Feedback -- 9.4.2 Event Sampled Feedback -- 9.4.3 Hybrid Reinforcement Learning Scheme -- 9.5 Perspectives on Controller Design with Image Feedback -- 9.6 Simulation Results -- 9.6.1 Linear Quadratic Regulator with Known Internal Dynamics -- 9.6.2 Optimal Adaptive Control with Unknown Drift Dynamics -- 9.7 Conclusion -- References.

10 Dissipativity-Based Verification for Autonomous Systems in Adversarial Environments -- 10.1 Introduction -- 10.1.1 Related Work -- 10.1.2 Contributions -- 10.1.3 Structure -- 10.1.4 Notation -- 10.2 Problem Formulation -- 10.2.1 (Q,S,R)-Dissipative and L2-Gain Stable Systems -- 10.3 Learning-Based Distributed Cascade Interconnection -- 10.4 Learning-Based L2-Gain Composition -- 10.4.1 Q-Learning for L2-Gain Verification -- 10.4.2 L2-Gain Model-Free Composition -- 10.5 Learning-Based Lossless Composition -- 10.6 Discussion -- 10.7 Conclusion and Future Work -- References -- 11 Reinforcement Learning-Based Model Reduction for Partial Differential Equations: Application to the Burgers Equation -- 11.1 Introduction -- 11.2 Basic Notation and Definitions -- 11.3 RL-Based Model Reduction of PDEs -- 11.3.1 Reduced-Order PDE Approximation -- 11.3.2 Proper Orthogonal Decomposition for ROMs -- 11.3.3 Closure Models for ROM Stabilization -- 11.3.4 Main Result: RL-Based Closure Model -- 11.4 Extremum Seeking Based Closure Model Auto-Tuning -- 11.5 The Case of the Burgers Equation -- 11.6 Conclusion -- References -- Part III Multi-agent Systems and RL -- 12 Multi-Agent Reinforcement Learning: A Selective Overview of Theories and Algorithms -- 12.1 Introduction -- 12.2 Background -- 12.2.1 Single-Agent RL -- 12.2.2 Multi-Agent RL Framework -- 12.3 Challenges in MARL Theory -- 12.3.1 Non-unique Learning Goals -- 12.3.2 Non-stationarity -- 12.3.3 Scalability Issue -- 12.3.4 Various Information Structures -- 12.4 MARL Algorithms with Theory -- 12.4.1 Cooperative Setting -- 12.4.2 Competitive Setting -- 12.4.3 Mixed Setting -- 12.5 Application Highlights -- 12.5.1 Cooperative Setting -- 12.5.2 Competitive Setting -- 12.5.3 Mixed Settings -- 12.6 Conclusions and Future Directions -- References.

13 Computational Intelligence in Uncertainty Quantification for Learning Control and Differential Games -- 13.1 Introduction -- 13.2 Problem Formulation of Optimal Control for Uncertain Systems -- 13.2.1 Optimal Control for Systems with Parameters Modulated by Multi-dimensional Uncertainties -- 13.2.2 Optimal Control for Random Switching Systems -- 13.3 Effective Uncertainty Evaluation Methods -- 13.3.1 Problem Formulation -- 13.3.2 The MPCM -- 13.3.3 The MPCM-OFFD -- 13.4 Optimal Control Solutions for Systems with Parameter Modulated by Multi-dimensional Uncertainties -- 13.4.1 Reinforcement Learning-Based Stochastic Optimal Control -- 13.4.2 Q-Learning-Based Stochastic Optimal Control -- 13.5 Optimal Control Solutions for Random Switching Systems -- 13.5.1 Optimal Controller for Random Switching Systems -- 13.5.2 Effective Estimator for Random Switching Systems -- 13.6 Differential Games for Systems with Parameters Modulated by Multi-dimensional Uncertainties -- 13.6.1 Stochastic Two-Player Zero-Sum Game -- 13.6.2 Multi-player Nonzero-Sum Game -- 13.7 Applications -- 13.7.1 Traffic Flow Management Under Uncertain Weather -- 13.7.2 Learning Control for Aerial Communication Using Directional Antennas (ACDA) Systems -- 13.8 Summary -- References -- 14 A Top-Down Approach to Attain

Decentralized Multi-agents -- 14.1 Introduction -- 14.2 Background -- 14.2.1 Reinforcement Learning -- 14.2.2 Multi-agent Reinforcement Learning -- 14.3 Centralized Learning, But Decentralized Execution -- 14.3.1 A Bottom-Up Approach -- 14.3.2 A Top-Down Approach -- 14.4 Centralized Expert Supervises Multi-agents -- 14.4.1 Imitation Learning -- 14.4.2 CESMA -- 14.5 Experiments -- 14.5.1 Decentralization Can Achieve Centralized Optimality -- 14.5.2 Expert Trajectories Versus Multi-agent Trajectories -- 14.6 Conclusion -- References.

15 Modeling and Mitigating Link-Flooding Distributed Denial-of-Service Attacks via Learning in Stackelberg Games.
