Preface

With the rapid development in information science and technology, many businesses and industries have undergone great changes, such as chemical industry, electric power engineering, electronics industry, mechanical engineering, transportation, and logistics business. While the scale of industrial enterprises is increasing, production equipment and industrial processes are becoming more and more complex. For these complex systems, decision and control are necessary to ensure that they perform properly and meet prescribed performance objectives. Under this circumstance, how to design safe, reliable, and efficient control for complex systems is essential for our society. As modern systems become more complex and performance requirements become more stringent, advanced control methods are greatly needed to achieve guaranteed performance and satisfactory goals.

In general, optimal control deals with the problem of finding a control law for a given system such that a certain optimality criterion is achieved. The main difference between optimal control of linear and nonlinear systems lies in that the latter often requires solving the nonlinear Bellman equation instead of the Riccati equation. Although dynamic programming is a conventional method in solving optimization and optimal control problems, it often suffers from the "curse of dimensionality." To overcome this difficulty, based on function approximators such as neural networks, adaptive/approximate dynamic programming (ADP) was proposed by Werbos as a method for solving optimal control problems forward-in-time.

This book presents the recent results of ADP with applications in optimal control. It is composed of 14 chapters which cover most of the hot research areas of ADP and are divided into three parts. Part I concerns discrete-time systems, including five chapters from Chaps. 2 to 6. Part II concerns continuous-time systems, including five chapters from Chaps. 7 to 11. Part III concerns applications, including three chapters from Chaps. 12 to 14.

In Chap. 1, an introduction to the history of ADP is provided, including the basic and iterative forms of ADP. The review begins with the origin of ADP and describes the basic structures and the algorithm development in detail. Connections between ADP and reinforcement learning are also discussed.

Part I: Discrete-Time Systems (Chaps. 2–6)

In Chap. 2, optimal control problems of discrete-time nonlinear dynamical systems, including optimal regulation, optimal tracking control, and constrained optimal control, are studied using a series of value iteration ADP approaches. First, an ADP scheme based on general value iteration is developed to obtain near-optimal control for discrete-time affine nonlinear systems with continuous state and control spaces. The present scheme is also employed to solve infinite-horizon optimal tracking control problems for a class of discrete-time nonlinear systems. In particular, using the globalized dual heuristic programming technique, a value iteration-based optimal control strategy of unknown discrete-time nonlinear dynamical systems with input constraints is established as a case study. Second, an iterative θ-ADP algorithm is given to solve the optimal control problem of infinite-horizon discrete-time nonlinear systems, which shows that each of the iterative controls can stabilize the nonlinear dynamical systems and the condition of initial admissible control is avoided effectively.

In Chap. 3, a series of iterative ADP algorithms are developed to solve the infinite-horizon optimal control problems for discrete-time nonlinear dynamical systems with finite approximation errors. Iterative control laws are obtained by using the present algorithms such that the iterative value functions reach the optimum. Then, the numerical optimal control problems are solved by a novel numerical adaptive learning control scheme based on ADP algorithm. Moreover, a general value iteration algorithm with finite approximate errors is developed to guarantee the iterative value function to converge to the solution of the Bellman equation. The general value iteration algorithm permits an arbitrary positive semidefinite function to initialize itself, which overcomes the disadvantage of traditional value iteration algorithms.

In Chap. 4, a discrete-time policy iteration ADP method is developed to solve the infinite-horizon optimal control problems for nonlinear dynamical systems. The idea is to use an iterative ADP technique to obtain iterative control laws that optimize the iterative value functions. The convergence, stability, and optimality properties are analyzed for policy iteration method for discrete-time nonlinear dynamical systems, and it is shown that the iterative value functions are nonincreasingly convergent to the optimal solution of the Bellman equation. It is also proven that any of the iterative control laws obtained from the present policy iteration algorithm can stabilize the nonlinear dynamical systems.

In Chap. 5, a generalized policy iteration algorithm is developed to solve the optimal control problems for infinite-horizon discrete-time nonlinear systems. Generalized policy iteration algorithm uses the idea of interacting the policy iteration algorithm and the value iteration algorithm of ADP. It permits an arbitrary positive semidefinite function to initialize the algorithm, where two iteration indices are used for policy evaluation and policy improvement, respectively. The monotonicity, convergence, admissibility, and optimality properties of the generalized policy iteration algorithm are analyzed.

In Chap. 6, error bounds of ADP algorithms are established for solving undiscounted infinite-horizon optimal control problems of discrete-time deterministic nonlinear systems. The error bounds for approximate value iteration based on a novel error condition are developed. The error bounds for approximate policy iteration and approximate optimistic policy iteration algorithms are also provided. It is shown that the iterative approximate value function can converge to a finite neighborhood of the optimal value function under some conditions. In addition, error bounds are also established for Q-function of approximate policy iteration for optimal control of unknown discounted discrete-time nonlinear systems. Neural networks are used to approximate the Q-function and the control policy.

Part II: Continuous-Time Systems (Chaps. 7–11)

In Chap. 7, optimal control problems of continuous-time affine nonlinear dynamical systems are studied using ADP approaches. First, an identifier–critic architecture based on ADP methods is presented to derive the approximate optimal control for uncertain continuous-time nonlinear dynamical systems. The identifier neural network and the critic neural network are tuned simultaneously, while the restrictive persistence of excitation condition is relaxed. Second, an ADP-based algorithm is developed to solve the optimal control problems for continuous-time nonlinear dynamical systems with control constraints. Only a single critic neural network is utilized to derive the optimal control, and there is no special requirement on the initial control.

In Chap. 8, the optimal control problems are considered for continuous-time nonaffine nonlinear dynamical systems with completely unknown dynamics via ADP methods. First, an ADP-based novel identifier–actor–critic architecture is developed to provide approximate optimal control solutions for continuous-time unknown nonaffine nonlinear dynamical systems, where the identifier is constructed by a dynamic neural network to transform nonaffine nonlinear systems into a class of affine nonlinear systems. Second, an ADP-based observer–critic architecture is presented to obtain the approximate optimal control for nonaffine nonlinear dynamical systems in the presence of unknown dynamics, where the observer is composed of a three-layer feedforward neural network aiming to get the knowledge of system states.

In Chap. 9, robust control and optimal guaranteed cost control of continuous-time uncertain nonlinear systems are studied using the idea of ADP. First, a novel strategy is established to design the robust controller for a class of nonlinear systems with uncertainties based on an online policy iteration algorithm. By properly choosing a cost function that reflects the uncertainties, regulation, and control, the robust control problem is transformed into an optimal control problem, which can be solved effectively under the framework of ADP. Then, the infinite-horizon optimal guaranteed cost control problem of uncertain nonlinear systems is investigated by employing the formulation of ADP-based online optimal control design, which extends the application scope of ADP methods to nonlinear and uncertain environment.

In Chap. 10, by using neural network-based online learning optimal control approach, a decentralized control strategy is developed to stabilize a class of continuous-time large-scale interconnected nonlinear systems. The decentralized control strategy of the overall system can be established by adding appropriate feedback gains to the optimal control laws of isolated subsystems. Then, an online policy iteration algorithm is presented to solve the Hamilton–Jacobi–Bellman equations related to the optimal control problems. Furthermore, as a generalization, a neural network-based decentralized control law is developed to stabilize the large-scale interconnected nonlinear systems with unknown dynamics by using an online model-free integral policy iteration algorithm.

In Chap. 11, differential game problems of continuous-time systems, including two-player zero-sum games, multiplayer zero-sum games, and multiplayer nonzero-sum games, are studied via a series of ADP approaches. First, an integral policy iteration algorithm is developed to learn online the Nash equilibrium solution of two-player zero-sum differential games with completely unknown continuous-time linear dynamics. Second, multiplayer zero-sum differential games for a class of continuous-time uncertain nonlinear systems are solved by using an iterative ADP algorithm. Finally, an online synchronous approximate optimal learning algorithm based on policy iteration is developed to solve multiplayer nonzero-sum games of continuous-time nonlinear systems without requiring exact knowledge of system dynamics.

Part III: Applications (Chaps. 12–14)

In Chap. 12, intelligent optimization methods based on ADP are applied to the challenges of intelligent price-responsive management of residential energy, with an emphasis on home battery use connected to the power grid. First, an action-dependent heuristic dynamic programming is developed to obtain the optimal control law for residential energy management. Second, a dual iterative Q-learning algorithm is developed to solve the optimal battery management and control problem in smart residential environments where two iterations are introduced, which are respectively internal and external iterations. Based on the dual iterative Q-learning algorithm, the convergence property of iterative Q-learning method for the optimal battery management and control problem is proven. Finally, a distributed iterative ADP method is developed to solve the multibattery optimal coordination control problem for home energy management systems.

In Chap. 13, a coal gasification optimal tracking control problem is solved through a data-based iterative optimal learning control scheme by using iterative ADP approach. According to system data, neural networks are used to construct the dynamics of coal gasification process, coal quality, and reference control, respectively. Via system transformation, the optimal tracking control problem with approximation errors and disturbances is effectively transformed into a two-person zero-sum optimal control problem. An iterative ADP algorithm is developed to obtain the optimal control laws for the transformed system.

In Chap. 14, a data-driven stable iterative ADP algorithm is developed to solve the optimal temperature control problem of water gas shift reaction system. According to the system data, neural networks are used to construct the dynamics of water gas shift reaction system and solve the reference control. Considering the reconstruction errors of neural networks and the disturbances of the system and control input, a stable iterative ADP algorithm is developed to obtain the optimal control law. Convergence property is developed to guarantee that the iterative value function converges to a finite neighborhood of the optimal cost function. Stability property is developed so that each of the iterative control laws can guarantee the tracking error to be uniformly ultimately bounded.

Beijing, China
Chicago, USA
September 2016

Derong Liu
Qinglai Wei
Ding Wang
Xiong Yang
Hongliang Li

Springer
http://www.springer.com/978-3-319-50813-9

Adaptive Dynamic Programming with Applications in Optimal Control

Liu, D.; Wei, Q.; Wang, D.; Yang, X.; Li, H.

2017, XXX, 594 p. 203 illus., 175 illus. in color., Hardcover
ISBN: 978-3-319-50813-9