Preface

Background of This Book

Optimal control, once thought of as one of the principal and complex domains in the control field, has been studied extensively in both science and engineering for several decades. As is known, dynamical systems are ubiquitous in nature and there exist many methods to design stable controllers for dynamical systems. However, stability is only a bare minimum requirement in the design of a system. Ensuring optimality guarantees the stability of nonlinear systems. As an extension of the calculus of variations, optimal control theory is a mathematical optimization method for deriving control policies. Dynamic programming is a very useful tool in solving optimization and optimal control problems by employing the principle of optimality. However, it is often computationally untenable to run true dynamic programming due to the well-known "curse of dimensionality". Hence, the adaptive dynamic programming (ADP) method was first proposed by Werbos in 1977. By building a system, called "critic", to approximate the cost function in dynamic programming, one can obtain the approximate optimal control solution to dynamic programming. In recent years, ADP algorithms have gained much attention from researchers in control fields. However, with the development of ADP algorithms, more and more people want to know the answers to the following questions:

(1) Are ADP algorithms convergent?
(2) Can the algorithm stabilize a nonlinear plant?
(3) Can the algorithm be run on-line?
(4) Can the algorithm be implemented in a finite time horizon?
(5) If the answer to the first question is positive, the subsequent questions are where the algorithm converges to, and how large the error is.

Before ADP algorithms can be applied to real plants, these questions need to be answered first. Throughout this book, we will study all these questions and give specific answers to each question.

Why This Book?

Although lots of monographs on ADP have appeared, the present book has unique features, which distinguish it from others.

First, the types of system involved in this monograph are rather extensive. From the point of view of models, one can find affine nonlinear systems, non-affine nonlinear systems, switched nonlinear systems, singularly perturbed systems and time-delay nonlinear systems in this book; these are the main mathematical models in the control fields.

Second, since the monograph is a summary of recent research works of the authors, the methods presented here for stabilizing, tracking, and games, which to a great degree benefit from optimal control theory, are more advanced than those appearing in introductory books. For example, the dual heuristic programming method is used to stabilize a constrained nonlinear system, with convergence proof; a data-based robust approximate optimal controller is designed based on simultaneous weight updating of two networks; and a single network scheme is proposed to solve the non-zero-sum game for a class of continuous-time systems.

Last but not least, some rather unique contributions are included in this monograph. One notable feature is the implementation of finite horizon optimal control for discrete-time nonlinear systems, which can obtain suboptimal control solutions within a fixed finite number of control steps. Most existing results in other books discuss only the infinite horizon control, which is not preferred in real-world applications. Besides this feature, another notable feature is that a pair of mixed optimal policies is developed to solve nonlinear games for the first time when the saddle point does not exist. Meanwhile, for the situation that the saddle point exists, existence conditions of the saddle point are avoided.

The Content of This Book

The book involves ten chapters. As implied by the book title, the main content of the book is composed of three parts; that is, optimal feedback control, nonlinear games, and related applications of ADP. In the part on optimal feedback control, the edge-cutting results on ADP-based infinite horizon and finite horizon feedback control, including stabilization control, and tracking control are presented in a systematic manner. In the part on nonlinear games, both zero-sum game and non-zero-sum games are studied. For the zero-sum game, it is proved for the first time that the iterative policies converge to the mixed optimal solutions when the saddle point does not exist. For the non-zero-sum game, a single network is proposed to seek the Nash equilibrium for the first time. In the part of applications, a self-learning call admission control scheme is proposed for CDMA cellular networks, and meanwhile an engine torque and air-fuel ratio control scheme is studied in detail, based on ADP.

In Chap. 1, a brief introduction to the background and development of ADP is provided. The review begins with the origin of ADP, and the basic structures and algorithm development are narrated in chronological order. After that, we turn attention to control problems based on ADP. We present this subject regarding two aspects: feedback control based on ADP and nonlinear games based on ADP. We mention a few iterative algorithms from recent literature and point out some open problems in each case.

In Chap. 2, the optimal state feedback control problem is studied based on ADP for both infinite horizon and finite horizon. Three different structures of ADP are utilized to solve the optimal state feedback control strategies, respectively. First, considering a class of affine constrained systems, a new DHP method is developed to stabilize the system, with convergence proof. Then, due to the special advantages of GDHP structure, a new optimal control scheme is developed with discounted cost functional. Moreover, based on a least-square successive approximation method, a series of GHJB equations are solved to obtain the optimal control solutions. Finally, a novel finite-horizon optimal control scheme is developed to obtain the suboptimal control solutions within a fixed finite number of control steps. Compared with the existing results in the infinite-horizon case, the present finite-horizon optimal controller is preferred in real-world applications.

Chapter 3 presents some direct methods for solving the closed-loop optimal tracking control problem for discrete-time systems. Considering the fact that the performance index functions of optimal tracking control problems are quite different from those of optimal state feedback control problems, a new type of performance index function is defined. The methods are mainly based on iterative HDP and GDHP algorithms. We first study the optimal tracking control problem of affine nonlinear systems, and after that we study the optimal tracking control problem of non-affine nonlinear systems. It is noticed that most real-world systems need to be effectively controlled within a finite time horizon. Hence, based on the above results, we further study the finite-horizon optimal tracking control problem, using the ADP approach in the last part of Chap. 3.

In Chap. 4, the optimal state feedback control problems of nonlinear systems with time delays are studied. In general, the optimal control for time-delay systems is an infinite-dimensional control problem, which is very difficult to solve; there are presently no good methods for dealing with this problem. In this chapter, the optimal state feedback control problems of nonlinear systems with time delays both in states and controls are investigated. By introducing a delay matrix function, the explicit expression of the optimal control function can be obtained. Next, for nonlinear time-delay systems with saturating actuators, we further study the optimal control problem using a non-quadratic functional, where two optimization processes are developed for searching the optimal solutions. The above two results are for the infinite-horizon optimal control problem. To the best of our knowledge, there are no results on the finite-horizon optimal control of nonlinear time-delay systems. Hence, in the last part of this chapter, a novel optimal control strategy is developed to solve the finite-horizon optimal control problem for a class of time-delay systems.

In Chap. 5, the optimal tracking control problems of nonlinear systems with time delays are studied using the HDP algorithm. First, the HJB equation for discrete time-delay systems is derived based on state error and control error. Then, a novel iterative HDP algorithm containing the iterations of state, control law, and cost functional is developed. We also give the convergence proof for the present iterative HDP algorithm. Finally, two neural networks, i.e., the critic neural network and the action neural network, are used to approximate the value function and the corresponding control law, respectively. It is the first time that the optimal tracking control problem of nonlinear systems with time delays is solved using the HDP algorithm.

In Chap. 6, we focus on the design of controllers for continuous-time systems via the ADP approach. Although many ADP methods have been proposed for continuous-time systems, a suitable framework in which the optimal controller can be designed for a class of general unknown continuous-time systems still has not been developed. In the first part of this chapter, we develop a new scheme to design optimal robust tracking controllers for unknown general continuous-time nonlinear systems. The merit of the present method is that we require only the availability of input/output data, instead of an exact system model. The obtained control input can be guaranteed to be close to the optimal control input within a small bound. In the second part of the chapter, a novel ADP-based robust neural network controller is developed for a class of continuous-time non-affine nonlinear systems, which is the first attempt to extend the ADP approach to continuous-time non-affine nonlinear systems.

In Chap. 7, several special optimal feedback control schemes are investigated. In the first part, the optimal feedback control problem of affine nonlinear switched systems is studied. To seek optimal solutions, a novel two-stage ADP method is developed. The algorithm can be divided into two stages: first, for each possible mode, calculate the associated value function, and then select the optimal mode for each state. In the second and third parts, the near-optimal controllers for nonlinear descriptor systems and singularly perturbed systems are solved by iterative DHP and HDP algorithms, respectively. In the fourth part, the near-optimal state-feedback control problem of nonlinear constrained discrete-time systems is solved via a single network ADP algorithm. At each step of the iterative algorithm, a neural network is utilized to approximate the costate function, and then the optimal control policy of the system can be computed directly according to the costate function, which removes the action network appearing in the ordinary ADP structure.

Game theory is concerned with the study of decision making in a situation where two or more rational opponents are involved under conditions of conflicting interests. In Chap. 8, zero-sum games are investigated for discrete-time systems based on the model-free ADP method. First, an effective data-based optimal control scheme is developed via the iterative ADP algorithm to find the optimal controller of a class of discrete-time zero-sum games for Roesser type 2-D systems. Since the exact models of many 2-D systems cannot be obtained inherently, the iterative ADP method is expected to avoid the requirement of exact system models. Second, a data-based optimal output feedback controller is developed for solving the zero-sum games of a class of discrete-time systems, whose merit is that knowledge of the model of the system is not required, nor the information of system states.

In Chap. 9, nonlinear game problems are investigated for continuous-time systems, including infinite horizon zero-sum games, finite horizon zero-sum games and non-zero-sum games. First, for the situations that the saddle point exists, the ADP technique is used to obtain the optimal control pair iteratively. The present approach makes the performance index function reach the saddle point of the zero-sum differential games, while complex existence conditions of the saddle point are avoided. For the situations that the saddle point does not exist, the mixed optimal control pair is obtained to make the performance index function reach the mixed optimum. Then, finite horizon zero-sum games for a class of nonaffine nonlinear systems are studied. Moreover, besides the zero-sum games, the non-zero-sum differential games are studied based on single network ADP algorithm. For zero-sum differential games, two players work on a cost functional together and minimax it. However, for non-zero-sum games, the control objective is to find a set of policies that guarantee the stability of the system and minimize the individual performance function to yield a Nash equilibrium.

In Chap. 10, the optimal control problems of modern wireless networks and automotive engines are studied by using ADP methods. In the first part, a novel learning control architecture is proposed based on adaptive critic designs/ADP, with only a single module instead of two or three modules. The choice of utility function for the present self-learning control scheme makes the present learning process much more efficient than existing learning control methods. The call admission controller can perform learning in real time as well as in off-line environments, and the controller improves its performance as it gains more experience. In the second part, an ADP-based learning algorithm is designed according to certain criteria and calibrated for vehicle operation over the entire operating regime. The algorithm is optimized for the engine in terms of performance, fuel economy, and tailpipe emissions through a significant effort in research and development and calibration processes. After the controller has learned to provide optimal control signals under various operating conditions off-line or on-line, it is applied to perform the task of engine control in real time. The performance of the controller can be further refined and improved through continuous learning in real-time vehicle operations.

Acknowledgments

The authors would like to acknowledge the help and encouragement they received during the course of writing this book. A great deal of the materials presented in this book is based on the research that we conducted with several colleagues and former students, including Q.L. Wei, Y. Zhang, T. Huang, O. Kovalenko, L.L. Cui, X. Zhang, R.Z. Song and N. Cao. We wish to acknowledge especially Dr. J.L. Zhang and Dr. C.B. Qin for their hard work on this book. The authors also wish to thank Prof. R.E. Bellman, Prof. D.P. Bertsekas, Prof. F.L. Lewis, Prof. J. Si and Prof. S. Jagannathan for their excellent books on the theory of optimal control and adaptive dynamic programming. We are very grateful to the National Natural Science Foundation of China (50977008, 60904037, 61034005, 61034002, 61104010), the Science and Technology Research Program of The Education Department of Liaoning Province (LT2010040), which provided necessary financial support for writing this book.

Shenyang, China — Huaguang Zhang
Beijing, China — Derong Liu
Chicago, USA — Yanhong Luo
Ding Wang

http://www.springer.com/978-1-4471-4756-5

Adaptive Dynamic Programming for Control
Algorithms and Stability
Zhang, H.; Liu, D.; Luo, Y.; Wang, D.
2013, XVI, 424 p., Hardcover
ISBN: 978-1-4471-4756-5