

Guest Editorial

Special Issue on Adaptive Dynamic Programming and Reinforcement Learning in Feedback Control

WE are extremely pleased to present this special issue of the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, PART B. The importance of adaptive dynamic programming (ADP) to feedback control engineers is that it affords a methodology for learning *optimal* control actions *online* in real time based on system performance *without necessarily knowing* the system dynamics. When such knowledge is required in ADP, it may be of low fidelity.

In feedback control engineering, various types of adaptive controllers provide implementation strategies that employ on-line observations of system performance to determine regulation controllers that drive the system to the equilibrium state, or tracking controllers that cause the system to follow prescribed trajectories. Certain techniques have been developed for online controller tuning without knowing the system dynamics. However, the control engineer is often constrained in the choice of performance measure or cost employed for the optimization. For example, inverse optimal adaptive controllers exist that optimize some derived performance measures that are reasonable though not of the control engineer's choosing. Indirect optimal adaptive controllers have been developed that require high-fidelity identification of the system dynamics.

It is becoming more and more clear that ADP techniques, on the other hand, do allow the design of *optimal controllers* online in real time in terms of a (freely) prescribed performance measure. The key lies effectively in solving Bellman's optimality condition forward in time through repeated iterations that involve: 1) computing the cost or value of using a current control, then 2) based on that value performing a control policy update, or control improvement. This can be viewed as a type of "structured" reinforcement learning comprising two components, a critic agent and a policy-update agent. The former evaluates currently instantiated control (via procedures called policy iteration or value iteration), and the latter improves controller design based on the latest evaluation. Typically, to allow practical implementation, neural networks are used in these respective agents (fuzzy logic systems are also possible) for value function approximation in the one case and control policy

approximation in the other. In the linear-system quadratic-cost-function case, the critic neural network is quadratic in the system state, and the neural net weights are exactly the entries of the Riccati equation solution matrix.

ADP has two important roles for the control engineer, specifically as follows. Riccati equation design has shown itself the backbone of modern control systems theory for linear quadratic control, but solution of the corresponding Riccati equation requires full knowledge of the system dynamics. It is also done *a priori* offline. ADP, on the other hand, allows solution of this Riccati equation online without (full) knowledge of the system dynamics. Arguably more important, ADP *extends* Riccati-equation-like design methods to *nonlinear* systems by using neural networks, of paradigms that are known to be universal function approximators.

Many of the practitioners in ADP over the years are represented in this special issue, which is broadly divided into three sections: Theoretical Foundations, Theory/Applications, and Applications. We are privileged to have a foreword written by Paul Werbos, the founder of ADP. A lead-in paper by George Lendaris sets into perspective historical, recent, and perhaps future developments in ADP.

F. L. LEWIS, *Guest Editor*
Electrical Engineering Department
University of Texas
Arlington, TX 76019 USA

DERONG LIU, *Guest Editor*
Department of Electrical and Computer Engineering
University of Illinois at Chicago
Chicago, IL 60607 USA

GEORGE G. LENDARIS, *Guest Editor*
Systems Science Graduate Program
Department of Electrical & Computer Engineering
Portland State University
Portland, OR 97207 USA



F. L. Lewis (S'78–M'81–SM'86–F'94) received the Bachelor's degree in physics/electrical engineering and the M.S.E.E. degree from Rice University, Houston, TX, the M.S. degree in aeronautical engineering from the University of West Florida, Pensacola, and the Ph.D. degree from the Georgia Institute of Technology, Atlanta.

He is currently a Distinguished Scholar Professor and the Moncrief-O'Donnell Chair at the Automation and Robotics Research Institute, University of Texas at Arlington (UTA). He works in feedback control, intelligent systems, and sensor networks. He is author of five U.S. patents, 186 journal papers, 305 conference papers, and 12 books.

Dr. Lewis received the Fulbright Research Award, NSF Research Initiation Grant, and ASEE Terman Award. He received the Outstanding Service Award from the Dallas IEEE Section and was selected as Engineer of the year by the Ft. Worth IEEE Section. He is listed in the Ft. Worth Business Press Top 200 Leaders in Manufacturing. He was appointed to the NAE Committee on Space Station in 1995. He is an elected Guest Consulting Professor at both the South China

University of Technology and Shanghai Jiao Tong University. He is also a Founding Member of the Board of Governors of the Mediterranean Control Association. He helped win the IEEE Control Systems Society Best Chapter Award (as Founding Chairman of DFW Chapter), the National Sigma Xi Award for Outstanding Chapter (as President of UTA Chapter), and the U.S. SBA Tibbets Award in 1996 (as Director of ARRI's SBIR Program). He is a Fellow of the IFAC, a Fellow of the U.K. Institute of Measurement and Control, PE Texas, and a Chartered Engineer in the U.K.



Derong Liu (S'91–M'94–SM'96–F'05) received the Ph.D. degree in electrical engineering from the University of Notre Dame, Notre Dame, IN, in 1994.

He is currently a Full Professor of electrical and computer engineering with the Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, IL. He has published seven books (four research monographs and three edited volumes). He is also an Associate Editor for *Automatica*.

Dr. Liu is currently the Editor for the IEEE Computational Intelligence Society's *ELECTRONIC LETTER*, an Associate Editor for the *IEEE TRANSACTIONS ON NEURAL NETWORKS*, an Associate Editor for the *IEEE COMPUTATIONAL INTELLIGENCE MAGAZINE*, and an Associate Editor for the *IEEE CIRCUITS AND SYSTEMS MAGAZINE*. He is the General Chair for the 2008 IEEE International Conference on Networking, Sensing and Control in China and the Program Chair for the 2008 International Joint Conference on Neural Networks. He is an elected AdCom member (2006–2008) of the IEEE Computational Intelligence Society. He was the recipient of

the Michael J. Birck Fellowship from the University of Notre Dame (1990), the Harvey N. Davis Distinguished Teaching Award from the Stevens Institute of Technology (1997), the Faculty Early Career Development (CAREER) Award from the National Science Foundation (1999), and the University Scholar Award from the University of Illinois (2006–2009).



George G. Lendaris (S'55–M'58–SM'74–F'83–LF'97) received the B.S., M.S., and Ph.D. degrees from the University of California, Berkeley, in 1957, 1958, and 1961, respectively.

He is currently the Director of the NW Computational Intelligence Laboratory and the Director of the Systems Science Graduate Program, Portland State University (PSU), Portland, OR. After receiving the Ph.D. degree, he joined GM Defense Research Laboratories, where he did extensive work in control systems, neural networks, and pattern recognition. In 1969, he joined academia, first, at the Oregon Graduate Institute as a Faculty Chair and, two years later, at PSU as one of the founders and developers of its Systems Science Graduate Program, serving there for the past 36 years. While there, he expanded his academic and research activities into general system theory and practice and back again to computational intelligence, focusing, for the past ten years, on the development of the adaptive critic form of reinforcement learning and its application as an approximate dynamic programming methodology to optimal control systems design and, more recently, to context discernment and experience-based methods of optimal control. He has served

a number of capacities at PSU, including President of the Faculty Senate, and has been active in neural network professional activities since the early 1990s.

Dr. Lendaris was the General Chair of the 1993 IJCNN; a member of the IEEE Neural Networks Council and, its successor, the IEEE Computational Intelligence Society Adcom during 2001–2007; a member of the International Neural Network Society (INNS) Board of Governors during 1997–2000 and 2003–2009; President of INNS during 2000–2003; and Local Chair of the 2003 IJCNN. He developed the optical diffraction pattern sampling methods of pattern recognition and was declared "Father of Diffraction Pattern Sampling" by The International Society for Optical Engineers in 1977.

IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS

PART B: CYBERNETICS

A PUBLICATION OF THE IEEE SYSTEMS, MAN, AND CYBERNETICS SOCIETY

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



AUGUST 2008

VOLUME 38

NUMBER 4

ITSCFI

(ISSN 1083-4419)

SPECIAL ISSUE ON ADAPTIVE DYNAMIC PROGRAMMING AND REINFORCEMENT LEARNING IN FEEDBACK CONTROL

GUEST EDITORIAL

Special Issue on Adaptive Dynamic Programming and Reinforcement Learning in Feedback Control	
..... F. L. Lewis, D. Liu, and G. G. Lendaris	896
ADP: The Key Direction for Future Research in Intelligent Control and Understanding Brain Intelligence	
..... P. J. Werbos	898

SPECIAL ISSUE PAPERS

Higher Level Application of ADP: A Next Phase for the Control Field?	G. G. Lendaris	901
Issues on Stability of ADP Feedback Controllers for Dynamical Systems (<i>Invited Paper</i>)		
..... S. N. Balakrishnan, J. Ding, and F. L. Lewis		913
Theoretical Foundations		
Hamilton–Jacobi–Bellman Equations and Approximate Dynamic Programming on Time Scales		
..... J. Seiffertt, S. Sanyal, and D. C. Wunsch, II		918
Random Sampling of States in Dynamic Programming	C. G. Atkeson and B. J. Stephens	924
Ensemble Algorithms in Reinforcement Learning	M. A. Wiering and H. van Hasselt	930
A Novel Infinite-Time Optimal Tracking Control Scheme for a Class of Discrete-Time Nonlinear Systems via the Greedy HDP Iteration Algorithm	H. Zhang, Q. Wei, and Y. Luo	937
Discrete-Time Nonlinear HJB Solution Using Approximate Dynamic Programming: Convergence Proof		
..... A. Al-Tamimi, F. L. Lewis, and M. Abu-Khalaf		943
Reinforcement Learning in Continuous Time and Space: Interference and Not Ill Conditioning Is the Main Problem When Using Distributed Function Approximators	B. Baddeley	950
Incoherent Control of Quantum Systems With Wavefunction-Controllable Subspaces via Quantum Reinforcement Learning	D. Y. Dong, C. L. Chen, T.-J. Tarn, A. Pechen, and H. Rabitz	957

(Contents Continued on Page 895)

Theory/Applications	
An Evolutionary Approach Toward Dynamic Self-Generated Fuzzy Inference Systems	<i>Y. Zhou and M. J. Er</i> 963
Decentralized Bayesian Search Using Approximate Dynamic Programming Methods	<i>Y. Zhao, S. D. Patek, and P. A. Beling</i> 970
Decentralized Learning in Markov Games	<i>P. Vrancx, K. Verbeeck, and A. Nowé</i> 976
Adaptive Feedback Control by Constrained Approximate Dynamic Programming	<i>S. Ferrari, J. E. Steck, and R. Chandramohan</i> 982
Applications	
Adaptive Critic Learning Techniques for Engine Torque and Air–Fuel Ratio Control	<i>D. Liu, H. Javaherian, O. Kovalenko, and T. Huang</i> 988
Control of Nonaffine Nonlinear Discrete-Time Systems Using Reinforcement-Learning-Based Linearly Parameterized Neural Networks	<i>Q. Yang, J. B. Vance, and S. Jagannathan</i> 994
Comparison of Adaptive Critic-Based and Classical Wide-Area Controllers for Power Systems	<i>S. Ray, G. K. Venayagamoorthy, B. Chaudhuri, and R. Majumder</i> 1002
Direct Heuristic Dynamic Programming for Damping Oscillations in a Large Power System	<i>C. Lu, J. Si, and X. Xie</i> 1008
Improved Adaptive–Reinforcement Learning Control for Morphing Unmanned Air Vehicles	<i>J. Valasek, J. Doebller, M. D. Tandale, and A. J. Meade</i> 1014

REGULAR ISSUE PAPERS	
Improving Iris Recognition Performance Using Segmentation, Quality Enhancement, Match Score Fusion, and Indexing	<i>M. Vatsa, R. Singh, and A. Noore</i> 1021
Instruction-Matrix-Based Genetic Programming	<i>G. Li, J. F. Wang, K. H. Lee, and K.-S. Leung</i> 1036
Adaptive Lyapunov-Based Control of a Robot and Mass–Spring System Undergoing an Impact Collision	<i>K. Dupree, C.-H. Liang, G. Hu, and W. E. Dixon</i> 1050
A Multifaceted Perspective at Data Analysis: A Study in Collaborative Intelligent Agents	<i>W. Pedrycz and P. Rai</i> 1062
Global Synchronization Control of General Delayed Discrete-Time Networks With Stochastic Coupling and Disturbances	<i>J. Liang, Z. Wang, Y. Liu, and X. Liu</i> 1073
Stability Analysis of Swarms With General Topology	<i>W. Li</i> 1084
Uncertainty Modeling of Improved Fuzzy Functions With Evolutionary Systems	<i>A. Celikyilmaz and I. B. Turksen</i> 1098
Nonlinear Dimensionality Reduction of Data Lying on the Multicenter Manifold	<i>D. Meng, Y. Leung, T. Fung, and Z. Xu</i> 1111
A Novel Gaze Estimation System With One Calibration Point	<i>A. Villanueva and R. Cabeza</i> 1123
Estimating Object Proper Motion Using Optical Flow, Kinematics, and Depth Information	<i>J. Schmüdderich, V. Willert, J. Eggert, S. Rebhan, C. Goerick, G. Sagerer, and E. Körner</i> 1139

REGULAR ISSUE CORRESPONDENCE	
Exponential Stability Analysis for Neural Networks With Time-Varying Delay	<i>M. Wu, F. Liu, P. Shi, Y. He, and R. Yokoyama</i> 1152
Fast Diagnosis With Sensors of Uncertain Quality	<i>O. Erdinc, C. Brideau, P. Willett, and T. Kirubarajan</i> 1157
Dynamic Output Feedback Control Synthesis for Continuous-Time T–S Fuzzy Systems via a Switched Fuzzy Control Scheme	<i>J. Dong and G.-H. Yang</i> 1166
Binary Two-Dimensional PCA	<i>Y. Pang, D. Tao, Y. Yuan, and X. Li</i> 1176
