

# **2007 IEEE International Symposium on Approximate Dynamic Programming and Reinforcement Learning**

**Honolulu, HI  
1-5 April 2007**



**IEEE Catalog Number:**  
**ISBN:**

**07EX1572**  
**1-4244-0706-0**

# Table of Contents

<b>Fitted Q Iteration with CMACs</b> .....	1
<i>Stephan Timmer, Martin Riedmiller</i>	
<b>Reinforcement-Learning-based Magneto-Hydrodynamic Control of Hypersonic Flows</b> .....	9
<i>Nilesh V. Kulkarni, Minh Q. Phan</i>	
<b>A Novel Fuzzy Reinforcement Learning Approach in Two-Level Intelligent Control of 3-DOF Robot Manipulators</b> .....	20
<i>Nasser Sadati, Mohammad Mollaie Emamzadeh</i>	
<b>Knowledge Transfer using Local Features</b> .....	26
<i>Martin Stolle, Christopher G. Atkeson</i>	
<b>Particle Swarm Optimized Adaptive Dynamic Programming</b> .....	32
<i>Dongbin Zhao, Jianqiang Yi, Derong Liu</i>	
<b>Discrete-time nonlinear HJB solution using Approximate dynamic programming: Convergence Proof</b> .....	38
<i>Asma Al-Tamimi, Frank Lewis</i>	
<b>Dual Representations for Dynamic Programming and Reinforcement Learning</b> .....	44
<i>Tao Wang, Michael Bowling, Dale Schuurmans</i>	
<b>An Optimal ADP Algorithm for a High-Dimensional Stochastic Control Problem</b> .....	52
<i>Juliana Nascimento, Warren Powell</i>	
<b>Convergence of Model-Based Temporal Difference Learning for Control</b> .....	60
<i>Hado van Hasselt and Marco A. Wiering</i>	
<b>On a Successful Application of Multi-Agent Reinforcement Learning to Operations Research Benchmarks</b> .....	68
<i>Thomas Gabel and Martin Riedmiller</i>	
<b>The Effect of Bootstrapping in Multi-Automata Reinforcement Learning</b> .....	76
<i>Maarten Peeters, Katja Verbeeck and Ann Nowe</i>	
<b>A Theoretical Analysis of Cooperative Behavior in Multi-agent Q-learning</b> .....	84
<i>Ludo Waltman and Uzay Kaymak</i>	
<b>Hybrid Ant Colony Optimization Using Memetic Algorithm for Traveling Salesman Problem</b> .....	92
<i>Haibin Duan, Xiufen Yu</i>	
<b>Robust Dynamic Programming for Discounted Infinite-Horizon Markov Decision Processes with Uncertain Stationary Transition Matrices</b> .....	96
<i>Baohua Li, Jennie Si</i>	
<b>Model-Based Reinforcement Learning in Factored-State MDPs</b> .....	103
<i>Alexander L. Strehl</i>	
<b>Leader-Follower semi-Markov Decision Problems: Theoretical Framework and Approximate Solution</b> .....	111
<i>Kurian Tharakunnel, Siddhartha Bhattacharyya</i>	
<b>A Scalable Recurrent Neural Network Framework for Model-Free POMDPs</b> .....	119
<i>Zhenzhen Liu, Itamar Elhanany</i>	
<b>Reinforcement learning by backpropagation through an LSTM model/critic</b> .....	127
<i>Bram Bakker</i>	
<b>Discrete-Time Approximate Dynamic Programming using Wavelet Basis Function Neural Networks</b> .....	135
<i>Ning Jin, Derong Liu, Ting Huang and Zhongyu Pang</i>	
<b>The Knowledge Gradient Policy for Offline Learning with Independent Normal Rewards</b> .....	143
<i>Peter Frazier, Warren Powell</i>	
<b>A Recurrent Control Neural Network for Data Efficient Reinforcement Learning</b> .....	151
<i>Anton Maximilian Schaefer, Steffen Udluft and Hans-Georg Zimmermann</i>	
<b>Computing Optimal Stationary Policies for Multi-Objective Markov Decision Processes</b> .....	158
<i>Marco A. Wiering and Edwin D. de Jong</i>	

# Table of Contents

<b>Strategy Generation with Cognitive Distance in Two-Player Games</b> .....	166
<i>Kosuke Sekiyama, Ricardo Carnieri and Toshio Fukuda</i>	
<b>Identifying Trajectory Classes in Dynamic Tasks</b> .....	172
<i>Stuart O. Anderson, Siddhartha S. Srinivasa</i>	
<b>A Dynamic Programming Approach to Viability Problems</b> .....	178
<i>Pierre-Arnaud Coquelin, Sophie Martin, Remi Munos</i>	
<b>Randomly Sampling Actions In Dynamic Programming</b> .....	185
<i>Christopher G. Atkeson</i>	
<b>SVM Viability Controller Active Learning: Application to Bike Control</b> .....	193
<i>Laetitia Chapel, Guillaume Deffuant</i>	
<b>An Approximate Dynamic Programming Approach for Job Releasing and Sequencing in a Reentrant Manufacturing Line</b> .....	201
<i>José A. Ramírez-Hernández and Emmanuel Fernandez</i>	
<b>Using ADP to Understand and Replicate Brain Intelligence: the Next Level Design</b> .....	209
<i>Paul J. Werbos</i>	
<b>Opposition-Based Reinforcement Learning in the Management of Water Resources</b> .....	217
<i>M. Mahootchi, H. R. Tizhoosh, K. Ponnambalam</i>	
<b>Online Reinforcement Learning Neural Network Controller Design for Nanomanipulation</b> .....	225
<i>Qinmin Yang and S. Jagannathan</i>	
<b>Short-term Stock Market Timing Prediction under Reinforcement Learning Schemes</b> .....	233
<i>Hailin Li, Cihan H. Dagli and David Enke</i>	
<b>Dynamic optimization of the strength ratio during a terrestrial conflict</b> .....	241
<i>Alexandre Sztykgold, Gilles Coppin, Olivier Hudry</i>	
<b>Continuous-Time ADP for Linear Systems with Partially Unknown Dynamics</b> .....	247
<i>Draguna Vrabie, Murad Abu-Khalaf, Frank L. Lewis, Youyi Wang</i>	
<b>Evaluation of Policy Gradient Methods and Variants on the Cart-Pole Benchmark</b> .....	254
<i>Martin Riedmiller, Jan Peters, Stefan Schaal</i>	
<b>Using Reward-weighted Regression for Reinforcement Learning of Task Space Control</b> .....	262
<i>Jan Peters, Stefan Schaal</i>	
<b>Toward effective combination of off-line and on-line training in ADP framework</b> .....	268
<i>Danil Prokhorov</i>	
<b>Reinforcement Learning in Continuous Action Spaces</b> .....	272
<i>Hado van Hasselt and Marco A. Wiering</i>	
<b>Two Novel On-policy Reinforcement Learning Algorithms based on TD(lambda)-methods</b> .....	280
<i>Marco A. Wiering and Hado van Hasselt</i>	
<b>Opposition-Based Q(lambda) with Non-Markovian Update</b> .....	288
<i>Maryam Shokri, Hamid R. Tizhoosh, Mohamed S. Kamel</i>	
<b>Coordinated Reinforcement Learning for Decentralized Optimal Control</b> .....	296
<i>Daniel Yagan and Chen-Khong Tham</i>	
<b>An Approximate Dynamic Programming Strategy for Responsive Traffic Signal Control</b> .....	303
<i>Chen Cai</i>	
<b>DHP Adaptive Critic Motion Control of Autonomous Wheeled Mobile Robot</b> .....	311
<i>Wei-Song Lin and Ping-Chieh Yang</i>	
<b>Approximate Optimal Control-Based Neurocontroller with a State Observation System for Seedlings Growth in Greenhouse</b> .....	318
<i>H. D. Patiño, J. A. Pucheta, C. Schugurensky, R. Fullana and B. Kuchen</i>	

# Table of Contents

<b>Efficient Learning in Cellular Simultaneous Recurrent Neural Networks - The Case of Maze Navigation Problem.....</b>	<b>324</b>
<i>Roman Ilin, Robert Kozma, Paul J. Werbos</i>	
<b>Value-Iteration Based Fitted Policy Iteration: Learning with a Single Trajectory .....</b>	<b>330</b>
<i>Andras Antos, Csaba Szepesvari, Remi Munos</i>	
<b>Kernelizing LSPE(lambda).....</b>	<b>338</b>
<i>Tobias Jung, Daniel Polani</i>	
<b>Q-learning algorithm with continuous state spaces and finite decision set.....</b>	<b>346</b>
<i>KENGY BARTY, PIERRE GIRARDEAU, JEAN-SEBASTIEN ROY AND CYRILLE STRUGAREK</i>	
<b>Sparse Temporal Difference Learning using LASSO.....</b>	<b>352</b>
<i>Manuel Loth, Manuel Davy, Philippe Preux</i>	