2009 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning

(ADPRL)

Nashville, Tennessee, USA
30 March – 2 April 2009
Tutorial ADPRL-T: Adaptive Dynamic Programming and Reinforcement Learning for Control Applications
Tuesday, March 31, 8:30AM-10:30AM, Room: Two Rivers, Instructor: Frank Lewis, University of Texas at Arlington, USA

Session ADPRL-1: Keynote Lecture
Tuesday, March 31, 11:00AM-12:00PM, Room: Two Rivers, Chair: Derong Liu, Chinese Academy of Sciences, China

A Unified Framework for Temporal Difference Methods [#1019] ................................................................. 1
Dimitri Bertsekas
Massachusetts Institute of Technology, United States

Session ADPRL-2: Adaptive Dynamic Programming and Reinforcement Learning
Tuesday, March 31, 12:00PM-1:00PM, Room: Two Rivers, Chair: Derong Liu, Chinese Academy of Sciences, China

Efficient Data Reuse in Value Function Approximation [#1012] ................................................................. 8
Hirotaka Hachiya, Takayuki Akiyama, Masashi Sugiyama and Jan Peters
Department of Computer Science, Tokyo Institute of Technology, Japan; Max-Planck for Biological Cybernetics, Germany

Lili Cui, Huaguang Zhang, Derong Liu and Yongsu Kim
Information Science and Engineering, Northeastern, China; Department of Electrical and Computer Engineering, University of Illinois, United States; Information Science and Engineering, Northeastern, Korea, Republic of

ADHDP(lamda) Strategies Based Coordinated Ramps Metering with Queuing Consideration [#1024] ...... 22
Xuerui Bai, Dongbin Zhao and Jianqiang Yi
Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, China

Session ADPRL-3: ADP and RL for Controls
Tuesday, March 31, 2:00PM-4:00PM, Room: Two Rivers, Chair: Matthieu Geist, INRIA, France, and Jun Ma, Princeton University, USA

Algorithm and Stability of ATC Receding Horizon Control [#1010] ................................................................. 28
Hongwei Zhang, Jie Huang and Frank Lewis
The Chinese University of Hong Kong, Hong Kong; the University of Texas at Arlington, United States
Online Policy Iteration Based Algorithms to Solve the Continuous-Time Infinite Horizon Optimal Control Problem [1034] ...................................................................................................................................................... 36
Kyriakos Vamvoudakis, Draguna Vrabie and Frank Lewis
Automation and Robotics Research Institute, University of Texas at Arlington, United States

Real-Time Motor Control using Recurrent Neural Networks [1033] ............................................................... 42
Dongsung Huh and Emanuel Todorov
UCSD Neuroscience Department, United States; UCSD Cognitive Science Department, United States

Hierarchical Optimal Control of a 7-DOF Arm Model [1027] ............................................................... 50
Dan Liu and Emanuel Todorov
UCSD, United States

Coupling Perception and Action Using Minimax Optimal Control [1031] ........................................................... 58
Tom Erez and William Smart
Washington University in St. Louis, United States

Session ADPRL-4: Markov Decision Processes
Tuesday, March 31, 4:30PM-6:30PM, Room: Two Rivers, Chair: Eugene Feinberg, Stony Brook University, USA, and Draguna Vrabie, University of Texas Arlington, USA

A Convergent Recursive Least Squares Policy Iteration Algorithm for Multi-Dimensional Markov Decision Process with Continuous State and Action Spaces [1008] ............................................................... 66
Jun Ma and Warren Powell
Princeton University, United States

Basis Function Adaptation Methods for Cost Approximation in MDP [1038] ........................................................... 74
Huizhen Yu and Dimitri Bertsekas
University of Helsinki, Finland; Massachusetts Institute of Technology, United States

Executing Concurrent Actions with Multiple Markov Decision Processes [1030] ........................................................... 82
Elva Corona-Xelhuantzi, Eduardo Morales and Enrique Sucar
INAOE, Mexico

Iterative Local Dynamic Programming [1026] ...................................................................................................................................................... 90
Emanuel Todorov and Yuval Tassa
University of California San Diego, United States; Hebrew University Jerusalem, Israel
Adaptive Computation of Optimal Nonrandomized Policies in Constrained Average-Reward MDPs
[1036]....................................................................................................................................................................96
Eugene Feinberg
Stony Brook University, United States

Session ADPRL-5: Architecture of ADP and RL
Wednesday, April 1, 8:30AM-10:30AM, Room: Two Rivers, Chair: H. Daniel Patino, Universidad Nacional de San
Juan, Argentina, and Claudia Stern, University of Paderborn, Germany

The QV Family Compared to Other Reinforcement Learning Algorithms [1018] .................................101
Marco Wiering and Hado van Hasselt
University of Groningen, Netherlands; Utrecht University, Netherlands

Feature Discovery in Approximate Dynamic Programming [1017] .........................................................109
Philippe Preux, Sertan Girgin and Manuel Loth
INRIA, LIFL, Universite de Lille, France; INRIA, France

Inferring bounds on the performance of a control policy from a sample of trajectories [1011] ............117
Raphael Fonteneau, Susan Murphy, Louis Wehenkel and Damien Ernst
University of Liege, Belgium; University of Michigan, United States

Neural-Network-Based Reinforcement Learning Controller for Nonlinear Systems with Non-symmetric
Dead-zone Inputs [1002] .................................................................................................................................124
Xin Zhang, Huaguang Zhang, Derong Liu and Yongsu Kim
Information Science and Engineering, Northeastern University, China; Department of Electrical and Computer
Engineering, University of Illinois, United States; Information Science and Engineering, Northeastern University,
Korea, Republic of

Algorithms for Variance Reduction in a Policy-Gradient Based Actor-Critic Framework [1023] .........130
Yogesh Awate
(1) Indian Institute of Technology, Bombay (2) marketRx - A Cognizant Company, India

Session ADPRL-6: Policy Search in ADP and RL
Wednesday, April 1, 11:00AM-1:00PM, Room: Two Rivers, Chair: Damien Ernst, Belgian FNRS and University of
Liege, Belgium, and Jan Peters, MPI, Germany

The Knowledge Gradient Algorithm For Online Subset Selection [1004] ..................................................137
Ilya Ryzhov and Warren Powell
Princeton University, United States

Planning under Uncertainty, Ensembles of Disturbance Trees and Kernelized Discrete Action Spaces
[1006] ...............................................................................................................................................................145
Boris Defourny, Damien Ernst and Louis Wehenkel
University of Liege, Belgium
Policy Search with Cross-Entropy Optimization of Basis Functions [#1007] .......................................................... 153
Lucian Busoniu, Damien Ernst, Bart De Schutter and Robert Babuska
Delft University of Technology, Netherlands; Belgian FNRS and University of Liege, Belgium

Eigenfunction approximation methods for linearly-solvable optimal control problems[#1029] ............ 161
Emanuel Todorov
University of California San Diego, United States

Learning Continuous-Action Control Policies [#1025] ............................................................................................ 169
Jason Pazis and Michail Lagoudakis
Technical University of Crete, Greece

Session ADPRL-7: Statistical and Multiagent RL
Wednesday, April 1, 2:00PM-4:00PM, Room: Two Rivers, Chair: Marco Wiering, The Netherlands, and Evangelos Theodorou, University of Southern California, USA

A Theoretical and Empirical Analysis of Expected Sarsa [#1020] ................................................................. 177
Harm van Seijen, Hado van Hasselt, Shimon Whiteson and Marco Wiering
TNO Defense, Safety and Security, Netherlands; Utrecht University, Netherlands; University of Amsterdam, Netherlands; University of Groningen, Netherlands

Kalman Temporal Differences: the deterministic case [#1015] ........................................................................ 185
Matthieu Geist, Olivier Pietquin and Gabriel Fricout
Supelec / ArcelorMittal / Inria, France; Supelec, France; ArcelorMittal, France

Integrating Sporadic Imitation in Reinforcement Learning Robots [#1032] ................................................... 193
Willi Richert, Ulrich Scheller, Markus Koch, Bernd Kleinjohann and Claudius Stern
University of Paderborn, Germany

Bounds of Optimal Learning [#1035] .................................................................................................................. 199
Roman Belavkin
Middlesex University, United Kingdom

Multiagent reinforcement learning in extensive form games with complete information [#1016] .......... 205
Ali Akramizadeh, Mohammad Bagher Menhaj and Ahmad Afshar
Polytechnic University of Tehran, Iran

Session ADPRL-8: Applications of ADP and RL
Wednesday, April 1, 4:30PM-6:30PM, Room: Two Rivers, Chair: Emanuel Todorov, University of California - San Diego, USA, and H. Daniel Patino, Universidad Nacional de San Juan, Argentina
Practical numerical methods for stochastic optimal control of biological systems in continuous time and space [#1037] .......................................................... 212
Alex Simpkins and Emanuel Todorov
Department of Mechanical and Aerospace Engineering, University of California, San Diego, United States; Department of Cognitive Science, University of California, San Diego, United States

Path Integral-Based Stochastic Optimal Control for Rigid Body Dynamics [#1021] .......................... 219
Evangelos Theodorou, Jonas Buchli and Stefan Schaal
University of Southern California, United States

Using Reward-Weighted Imitation for Robot Reinforcement Learning [#1022] .......................... 226
Jan Peters and Jens Kober
Max Planck Institute for Biological Cybernetics, Germany

H. Daniel Patino, Santiago Tosetti and Flavio Capraro
Instituto de Automatica, Universidad Nacional de San Juan, Argentina

Neuro-controller of Cement Rotary Kiln Temperature with Adaptive Critic Designs [#1009] ............. 238
Xiaofeng Lin, Tangbo Liu, Shaojian Song and Chunning Song
College of Electrical Engineering of Guangxi University, China

AUTHOR INDEX