

2011 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning

(ADPRL 2011)

**Paris, France
11 – 15 April 2011**



**IEEE Catalog Number: CFP11ADP-PRT
ISBN: 978-1-4244-9887-1**

TABLE OF CONTENTS

ADPRL 2011 COMMITTEE.....	ix
ADPRL 2011 KEYNOTE	x
ADPRL 2011 TUTORIAL	xx
ADPRL 2011 TECHNICAL SESSIONS.....	1

Tuesday, April 12

09:30 - 10:30

S1: ADPRL - Keynote

Chair: Sarangapani Jagannathan (Missouri University of Science and Technology, USA)

Higher-Level Application of Adaptive Dynamic Programming/Reinforcement Learning - a Next Phase for Controls and System Identification?

George Lendaris (Portland State University, USA)

11:00 - 12:00

S2: Representations in Reinforcement Learning

Chairs: Michail G. Lagoudakis (Technical University of Crete, Greece), Emanuel Todorov (University of Washington, Seattle, USA)

Approximate Reinforcement Learning: An Overview..... 1

Lucian Busoniu (Delft University of Technology, The Netherlands)

Damien Ernst (University of Liège, Belgium)

Bart De Schutter (Delft University of Technology, The Netherlands)

Robert Babuska (Delft University of Technology, The Netherlands)

Parametric Value Function Approximation: a Unified View..... 9

Matthieu Geist (Supelec, France)

Olivier Pietquin (GeorgiaTech - CNRS, France)

On Learning with Imperfect Representations..... 17

Shivaram Kalyanakrishnan (The University of Texas at Austin, USA)

Peter Stone (University of Texas at Austin, USA)

14:00 – 16:00

S3: Active Reinforcement Learning

Chairs: Lucian Busoniu (Delft University of Technology, The Netherlands), Damien Ernst (University of Liège, Belgium)

Active Exploration for Robot Parameter Selection in Episodic Reinforcement Learning..... 25

Oliver Kroemer (Max Planck Institute, Germany)

Jan Peters (Max-Planck Institute, Germany)

Active Learning for Personalizing Treatment.....	32
Kun Deng (University of Michigan, USA)	
Joelle Pineau (McGill University, Canada)	
Susan Murphy (University of Michigan, USA)	
Active Exploration by Searching for Experiments that Falsify the Computed Control Policy.....	40
Raphael Fonteneau (University of Liège, Belgium)	
Susan Murphy (University of Michigan, USA)	
Louis Wehenkel (University of Liège, Belgium)	
Damien Ernst (University of Liège, Belgium)	
Optimistic Planning for Sparsely Stochastic Systems.....	48
Lucian Busoniu (Delft University of Technology, The Netherlands)	
Remi Munos (INRIA, France)	
Bart De Schutter (Delft University of Technology, The Netherlands)	
Robert Babuska (Delft University of Technology, The Netherlands)	
Adaptive Sample Collection Using Active Learning for Kernel-based Approximate Policy Iteration.....	56
Chunming Liu (National University of Defense Technology, China)	
Xin Xu (National University of Defense Technology, China)	
Haiyun Hu Xu (National University of Defense Technology, China)	
Bin Dai (National University of Defense Technology, China)	
Tree-based Variable Selection for Dimensionality Reduction of Large-scale Control Systems.....	62
Andrea Castelletti (Politecnico di Milano, Italy)	
Stefano Galelli (Politecnico di Milano, Italy)	
Marcello Restelli (Politecnico di Milano, Italy)	
Rodolfo Soncini-Sessa (Politecnico di Milano, Italy)	
 16:30 - 18:30	
S4: ADPRL - Tutorial	
ADP for Control of Continuous-Time Dynamical Systems, Multi-Player Differential Games, and Games on Graphs	
Frank Lewis (Automation and Robotics Research Institute, University of Texas at Arlington, USA)	
Draguna Vrabie (United Technologies Research Center, USA)	
 Wednesday, April 13	
08:30 - 10:30	
S5: Reinforcement Learning I – ADPRL	
Chairs: Damien Ernst (University of Liège, Belgium), Marco Wiering (University of Groningen, The Netherlands)	
High-Order Local Dynamic Programming.....	70
Yuval Tassa (Hebrew University, Israel)	
Emanuel Todorov (University of Washington, USA)	
Safe Reinforcement Learning in High-Risk Tasks through Policy Improvement.....	76
Francisco Javier Garcia Polo (Universidad Carlos III de Madrid, Spain)	
Fernando Fernandez Rebollo (Universidad Carlos III de Madrid, Spain)	

Agent Self-Assessment: Determining Policy Quality Without Execution	84
Alexander Hans (Ilmenau University of Technology, Germany)	
Siegmond Duell (Institute of Technology Berlin, Germany)	
Steffen Udluft (Siemens AG, Germany)	

Reinforcement Learning Algorithms for solving Classification Problems.....	91
Marco A. Wiering (University of Groningen, The Netherlands)	
Hado van Hasselt (Centrum Wiskunde & Informatica, The Netherlands)	
Auke-Dirk Pietersma (University of Groningen, The Netherlands)	
Lambert Schomaker (University of Groningen, The Netherlands)	

Reinforcement Learning in Multidimensional Continuous Action Spaces	97
Jason Papis (Duke University, USA)	
Michail G. Lagoudakis (Technical University of Crete, Greece)	

Grounding Subgoals in Information Transitions	105
Sander G. van Dijk (University of Hertfordshire, United Kingdom)	
Daniel Polani (University of Hertfordshire, United Kingdom)	

11:00 - 12:40

S6: Reinforcement Learning II

Chairs: Rajesh Ganesan (George Mason University, USA), Shimon Whiteson (University of Amsterdam, The Netherlands)

Directed Exploration of Policy Space using Support Vector Classifiers.....	112
Ioannis Rexakis (Technical University of Crete, Greece)	
Michail G. Lagoudakis (Technical University of Crete, Greece)	

Protecting Against Evaluation Overfitting in Empirical Reinforcement Learning.....	120
Shimon Whiteson (University of Amsterdam, The Netherlands)	
Brian Tanner (University of Alberta, Canada)	
Matthew E. Taylor (Lafayette College, USA)	
Peter Stone (University of Texas at Austin, USA)	

Higher Order Q-Learning	128
Ashley Edwards (University of Georgia, USA)	
William M. Pottenger (Rutgers University, USA)	

Adaptive Action Definition in Continuous Reinforcement Learning.....	B#5
Behrouz Haji Soleimani Mamaghani (University of Tehran, Iran);	
Mojid Nili-Ahmadabadi (University of Tehran, Iran)	
Babak Nadjar Araabi (University of Tehran, Iran)	

14:00 – 16:00

S7: Recent Advances in Reinforcement Learning

Chair: Abhijit Gosavi (Missouri S & T, USA)

Bayesian Active Learning With Basis Functions	143
Ilya O. Ryzhov (Princeton University, USA)	
Warren B. Powell (Princeton University, USA)	

Evolutionary Value Function Approximation	151
Mohsen Davarynejad (Delft University of Technology, The Netherlands)	
Jelmer van Ast (Delft Center for Systems and Control, The Netherlands)	
Jos Vrancken (Delft University of Technology, The Netherlands)	
Jan van den Berg (Delft University of Technology, The Netherlands)	
Enhancing the Episodic Natural Actor-Critic Algorithm by a Regularisation Term to Stabilize Learning of Control Structures	156
Andreas Witsch (Universitat Kassel, Germany)	
Roland Reichle (Universitat Kassel, Germany)	
Kurt Geihs (Universitat Kassel, Germany)	
Sascha Lange (Albert-Ludwigs-Universitat Freiburg, Germany)	
Martin Riedmiller (Albert-Ludwigs-Universitat Freiburg, Germany)	
Application of Reinforcement Learning-based Algorithms in CO₂ Allowance and Electricity Markets	164
Vishnuteja Nanduri (University of Wisconsin-Milwaukee, USA)	
Model-building Semi-Markov Adaptive Critics	170
Abhijit Gosavi (Missouri S & T, USA)	
Susan L. Murray (Missouri S & T, USA)	
Jiaqiao Hu (Stony Brook University, USA)	
Dynamic Lead Time Promising	176
Matthew J. Reindorp (Eindhoven University of Technology, The Netherlands)	
Michael C. Fu (University of Maryland, USA)	
 14:40 – 16:00	
S8: Approximate Dynamic Programming for Optimal Control of Nonlinear Systems; Approximate Dynamic Programming for Optimal Control of Nonlinear Systems	
Chair: Huaguang Zhang (Northeastern University, China)	
N-step optimal time-invariant trajectory tracking control for a class of nonlinear systems	184
Ruizhuo Song (Northeastern University, China)	
Huaguang Zhang (Northeastern University, China)	
Data-based adaptive critic design for discrete-time zero-sum games using output feedback.....	190
Lili Cui (Northeastern University, China)	
Huaguang Zhang (Northeastern University, China)	
Xin Zhang (Northeastern University, China)	
Yanhong Luo (Northeastern University, China)	
Global Optimal Strategies of a Class of Finite-horizon Continuous-time Nonaffine Nonlinear Zero-sum Game Using a New Iteration Algorithm.....	196
Xin Zhang (Northeastern University, China)	
Huaguang Zhang (Northeastern University, China)	
Lili Cui (Northeastern University, China)	
Yanhong Luo (Northeastern University, China)	
Path Integral Control and Bounded Rationality	202
Daniel A .Braun (University of Southern California, USA)	
Pedro A. Ortega (University of Cambridge, United Kingdom)	
Evangelos Theodorou (University of Southern California, USA)	
Stefan Schaal (University of Southern California, USA)	

Adaptive Dynamic Programming with Balanced Weights Seeking Strategy210
Jian Fu (University of Rhode Island, USA)
Haibo He (University of Rhode Island, USA)
Zhen Ni (University of Rhode Island, USA)

Moving Least-squares Approximations for Linearly-solvable MDP.....218
Mingyuan Zhong (University of Washington, USA)
Emanuel Todorov (University of Washington, Seattle, USA)

Taxi-out Time Analysis using Approximate Dynamic Programming B#5
Rajesh Ganesan (George Mason University, USA)
Lance Sherry (George Mason University, USA)
Poornima Balakrishna (Sensis Corporation, USA)

Thursday, April 14

8:30 – 10:30

S9: Approximate Dynamic Programming for Feedback Control; Multiagent Systems and Multiplayer Games

Chairs: Frank Lewis (Automation and Robotics Research Institute, University of Texas at Arlington, USA),
Draguna Vrabie (United Technologies Research Center, USA)

Near Optimal Control of Mobile Robot Formations234
Travis Dierks (DRS Sustainment Systems, Inc., USA)
Bryan Brenner (Missouri University of Science and Technology, USA)
S. Jagannathan (Missouri University of Science and Technology, USA)

Adaptive Dynamic Programming for Optimal Control of Unknown Nonlinear Discrete-Time Systems242
Derong Liu (Chinese Academy of Sciences, China)
Ding Wang (Chinese Academy of Sciences, China)
Dongbin Zhao (Chinese Academy of Sciences, China)

Online Adaptive Learning of Optimal Control Solutions Using Integral Reinforcement Learning250
Kyriakos Vamvoudakis (University of Texas at Arlington, USA)
Draguna Vrabie (University of Texas at Arlington, USA)
Frank L. Lewis (University of Texas at Arlington, University of Texas at Arlington, USA)

Online Near Optimal Control of Unknown Nonaffine systems with Application to HCCI Engines258
H. Zargarzadeh (Missouri University of Science and Technology, USA)
S. Jagannathan (Missouri University of Science and Technology, USA)
J. Drallmeier (Missouri University of Science and Technology, USA)

Policy gradient methods with model predictive control applied to ball bouncing B#5
Paul Kulchenko (University of Washington, USA)
Emanuel Todorov (University of Washington, USA)

An Approximate Dynamic Programming Based Controller for an Underactuated 6DoF Quadrotor271
Emanuel Stingu (University of Texas at Arlington, USA)
Frank L. Lewis (University of Texas at Arlington, USA)

11:00 – 12:00

S10: Applications of Reinforcement Learning

Chairs: Thomas Gabel (University of Freiburg, Germany), Martino Migliavacca (Politecnico di Milano, Italy)

Improved Neural Fitted Q Iteration Applied to a Novel Computer Gaming and Learning Benchmark.....279

Thomas Gabel (University of Freiburg, Germany)
Christian Lutz (University of Freiburg, Germany)
Martin Riedmiller (University of Freiburg, Germany)

Fitted Policy Search287

Martino Migliavacca (Politecnico di Milano, Italy)
Alessio Pecorino (Politecnico di Milano, Italy)
Matteo Pirota (Politecnico di Milano, Italy)
Marcello Restelli (Politecnico di Milano, Italy)
Andrea Bonarini (Politecnico di Milano, Italy)

A Reinforcement Learning Approach for Sequential Mastery Testing295

El-Sayed M. El-Alfy (King Fahd University of Petroleum and Minerals, Saudi Arabia)

14:00 – 16:00

S11: Optimal Control and Applications

Chairs: Hisashi Handa (Okayama University, Japan), Warren Powell (Princeton University, USA)

Information Space Receding Horizon Control.....302

Suman Chakravorty (Texas A&M University, USA);
R. Scott Erwin (Air Force Research Laboratory, USA)

Feedback Controller Parameterizations for Reinforcement Learning310

John W. Roberts (MIT, USA)
Ian R. Manchester (MIT, USA)
Russ Tedrake (MIT, USA)

Supervised Adaptive Dynamic Programming Based Adaptive Cruise Control318

Dongbin Zhao (Chinese Academy of Sciences, China)
Zhaohui Hu (Chinese Academy of Sciences, China)

An Adaptive-Learning Framework for Semi-Cooperative Multi-agent Coordination324

Abdeslem Boukhtouta (Defence Research and Development Canada-Valcartier, Canada)
Jean Berger (Defence Research and Development Canada-Valcartier, Canada)
Warren B. Powell (Princeton University, USA)
Abraham George (Princeton University, USA)

Structure Search of Probabilistic Models and Data Correction for EDA-RL332

Hisashi Handa (Okayama University, Japan)

Complex object manipulation with hierarchical optimal control.....338

Alex Simpkins (University of Washington, USA)
Emanuel Todorov (University of Washington, USA)

AUTHOR INDEX346