

Lecture Notes in Artificial Intelligence 7188

Subseries of Lecture Notes in Computer Science

LNAI Series Editors

Randy Goebel

University of Alberta, Edmonton, Canada

Yuzuru Tanaka

Hokkaido University, Sapporo, Japan

Wolfgang Wahlster

DFKI and Saarland University, Saarbrücken, Germany

LNAI Founding Series Editor

Joerg Siekmann

DFKI and Saarland University, Saarbrücken, Germany

Scott Sanner Marcus Hutter (Eds.)

Recent Advances in Reinforcement Learning

9th European Workshop, EWRL 2011
Athens, Greece, September 9-11, 2011
Revised Selected Papers

Series Editors

Randy Goebel, University of Alberta, Edmonton, Canada
Jörg Siekmann, University of Saarland, Saarbrücken, Germany
Wolfgang Wahlster, DFKI and University of Saarland, Saarbrücken, Germany

Volume Editors

Scott Sanner
NICTA and the Australian National University
7 London Circuit
Canberra, ACT 2601 Australia
E-mail: Scott.Sanner@nicta.com.au

Marcus Hutter
Australian National University
Research School of Computer Science
Canberra, ACT 0200 Australia
E-mail: marcus.hutter@anu.edu.au

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-29945-2 e-ISBN 978-3-642-29946-9
DOI 10.1007/978-3-642-29946-9
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012937733

CR Subject Classification (1998): I.2.6, I.2, F.1-2, G.3, C.1.3

LNCS Sublibrary: SL 7 – Artificial Intelligence

© Springer-Verlag Berlin Heidelberg 2012

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This volume contains the papers presented at EWRL 2011: the 9th European Workshop on Reinforcement Learning held in Athens, Greece, September 9–11, 2011. The workshop was co-located with the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2011). The technical program of EWRL 2011 contained 38 plenary talks by authors of submitted papers and 4 plenary invited talks:

- Peter Auer (University of Leoben, Austria), “UCRL and Autonomous Exploration”
- Kristian Kersting (Fraunhofer IAIS and University of Bonn, Germany), “Increasing Representational Power and Scaling Inference in Reinforcement Learning”
- Peter Stone (University of Texas Austin, USA), “PRISM Practical RL: Representation, Interaction, Synthesis, and Mortality”
- Csaba Szepesvári (University of Alberta, Canada), “Towards Robust Reinforcement Learning Algorithms”

All authors of presented papers were invited to present their papers at an evening poster session held on September 10; posters were also solicited for late-breaking results with the option to submit a full paper after the workshop.

In total, there were 40 submissions to the workshop from 58 authors representing 18 countries (9 European). All papers were reviewed by at least two reviewers, with some papers receiving a third review. We had an international Program Committee of 24 members and 9 additional reviewers residing in 11 countries (7 European). Reviewing proceeded in two rounds: 24 of 39 submitted papers were accepted to the LNAI proceedings in the first round; after the workshop authors were allowed to resubmit previous submissions or submit new papers from late-breaking posters in a second round of reviewing in which 4 of 14 submissions (13 resubmissions, 1 new submission) were accepted for the LNAI proceedings. The present volume contains the texts of the 28 papers accepted for the LNAI proceedings in the first and second round.

The best paper prize of 500 Euro was awarded to the paper “Automatic Discovery of Ranking Formulas for Playing with Multi-armed Bandits” by authors Francis Maes, Louis Wehenkel and Damien Ernst (University of Liège, Belgium).

EWRL was first held in 1995 and has been held every 2-3 years since that time. Over the years, EWRL has become one of the premier events for the discussion and dissemination of recent research results in the field of reinforcement learning. This year’s workshop was organized by Marcus Hutter (Australian National University, Canberra, Australia), Matthew Robards (Australian National University, Canberra, Australia), Scott Santer (NICTA and Australian National University, Canberra, Australia), Peter Sunehag (Australian National University,

Canberra, Australia), and Marco Weiring (University of Groningen, Groningen, The Netherlands).

We would like to thank our financial sponsors: the *Artificial Intelligence Journal*, the Australian National University, NICTA, and PASCAL2. Their generous support has allowed us to continue the tradition of holding EWRL as a free event open to all interested participants. We would like to thank ECML PKDD 2011 for allowing us to co-locate with them and for advertising EWRL on their website. We are grateful that we could use the excellent conference management system EasyChair, which has been developed mainly by Andrei Voronkov and hosted at the University of Manchester; the system is cost-free. We thank Springer for their support in preparing and publishing this volume of *Lecture Notes in Artificial Intelligence*.

Last but not least, we wish to thank the organizers, Program Committee members, additional reviewers, authors, speakers and all attendees for this engaging and highly successful installment of EWRL.

February 2012

Scott Sanner
Marcus Hutter

Organization

Organizing Committee

General Workshop Chair	Marcus Hutter
Local Organizing Chair	Matthew Robards
Program Committee Chair	Scott Sanner
Treasurer	Peter Sunehag
Miscellaneous	Marco Wiering

Program Committee

Edwin Bonilla	NICTA and Australian National University, Australia
Emma Brunskill	University of California, Berkeley, USA
Peter Dayan	University College London, UK
Carlos Diuk	Princeton University, USA
Alan Fern	Oregon State University, USA
Fernando Fernandez	Universidad Carlos III de Madrid, Spain
Mohammad Ghavamzadeh	INRIA Lille, France
Marcus Hutter	Australian National University and NICTA, Australia
Kristian Kersting	Fraunhofer IAIS and University of Bonn, Germany
Shie Mannor	Technion, Israel
Ronald Ortner	University of Leoben, Austria
Joelle Pineau	McGill University, Canada
Doina Precup	McGill University, Canada
Matthew Robards	Australian National University and NICTA, Australia
Scott Sanner	NICTA and Australian National University, Australia
Guy Shani	Ben-Gurion University, Israel
David Silver	University College London, UK
Peter Sunehag	Australian National University, Australia
Prasad Tadepalli	Oregon State University, USA
Matthew Taylor	Lafayette College, USA

VIII Organization

William Uther

NICTA and the University of New South Wales,
Australia

Martijn Van Otterlo

Katholieke Universiteit Leuven, Belgium

Thomas Walsh

University of Arizona, USA

Marco Wiering

University of Groningen, The Netherlands

Additional Reviewers

Bou Ammar, Haitham

Shào, Wén

Daswani, Mayank

Van Hasselt, Hado

Kalyanakrishnan, Shivaram

Vroman, Monica

Lattimore, Tor

Visentin, Daniel

Nguyen, Phuong Minh

With Thanks to Our EWRL 2011 Sponsors



Australian
National
University



PASCAL2

Pattern Analysis, Statistical Modelling and
Computational Learning

Table of Contents

Invited Talk Abstracts

Invited Talk: UCRL and Autonomous Exploration	1
<i>Peter Auer</i>	
Invited Talk: Increasing Representational Power and Scaling Inference in Reinforcement Learning	2
<i>Kristian Kersting</i>	
Invited Talk: PRISM – Practical RL: Representation, Interaction, Synthesis, and Mortality	3
<i>Peter Stone</i>	
Invited Talk: Towards Robust Reinforcement Learning Algorithms	4
<i>Csaba Szepesvári</i>	

Online Reinforcement Learning

Automatic Discovery of Ranking Formulas for Playing with Multi-armed Bandits	5
<i>Francis Maes, Louis Wehenkel, and Damien Ernst</i>	
Goal-Directed Online Learning of Predictive Models	18
<i>Sylvie C.W. Ong, Yuri Grinberg, and Joelle Pineau</i>	
Gradient Based Algorithms with Loss Functions and Kernels for Improved On-Policy Control	30
<i>Matthew Robards and Peter Sunehag</i>	

Learning and Exploring MDPs

Active Learning of MDP Models	42
<i>Mauricio Araya-López, Olivier Buffet, Vincent Thomas, and François Charpillet</i>	
Handling Ambiguous Effects in Action Learning	54
<i>Boris Lesner and Bruno Zanuttini</i>	
Feature Reinforcement Learning in Practice	66
<i>Phuong Nguyen, Peter Sunehag, and Marcus Hutter</i>	

Function Approximation Methods for Reinforcement Learning

Reinforcement Learning with a Bilinear Q Function	78
<i>Charles Elkan</i>	
ℓ_1 -Penalized Projected Bellman Residual	89
<i>Matthieu Geist and Bruno Scherrer</i>	
Regularized Least Squares Temporal Difference Learning with Nested ℓ_2 and ℓ_1 Penalization	102
<i>Matthew W. Hoffman, Alessandro Lazaric, Mohammad Ghavamzadeh, and Rémi Munos</i>	
Recursive Least-Squares Learning with Eligibility Traces	115
<i>Bruno Scherrer and Matthieu Geist</i>	
Value Function Approximation through Sparse Bayesian Modeling	128
<i>Nikolaos Tziortziotis and Konstantinos Blekas</i>	

Macro-actions in Reinforcement Learning

Automatic Construction of Temporally Extended Actions for MDPs Using Bisimulation Metrics	140
<i>Pablo Samuel Castro and Doina Precup</i>	
Unified Inter and Intra Options Learning Using Policy Gradient Methods	153
<i>Kfir Y. Levy and Nahum Shimkin</i>	
Options with Exceptions	165
<i>Munu Sairamesh and Balaraman Ravindran</i>	

Policy Search and Bounds

Robust Bayesian Reinforcement Learning through Tight Lower Bounds	177
<i>Christos Dimitrakakis</i>	
Optimized Look-ahead Tree Search Policies	189
<i>Francis Maes, Louis Wehenkel, and Damien Ernst</i>	
A Framework for Computing Bounds for the Return of a Policy	201
<i>Cosmin Păduraru, Doina Precup, and Joelle Pineau</i>	

Multi-Task and Transfer Reinforcement Learning

Transferring Evolved Reservoir Features in Reinforcement Learning Tasks	213
<i>Kyriakos C. Chatzidimitriou, Ioannis Partalas, Pericles A. Mitkas, and Ioannis Vlahavas</i>	
Transfer Learning via Multiple Inter-task Mappings	225
<i>Anestis Fachantidis, Ioannis Partalas, Matthew E. Taylor, and Ioannis Vlahavas</i>	
Multi-Task Reinforcement Learning: Shaping and Feature Selection	237
<i>Matthijs Snel and Shimon Whiteson</i>	

Multi-Agent Reinforcement Learning

Transfer Learning in Multi-Agent Reinforcement Learning Domains	249
<i>Georgios Boutsioukis, Ioannis Partalas, and Ioannis Vlahavas</i>	
An Extension of a Hierarchical Reinforcement Learning Algorithm for Multiagent Settings	261
<i>Ioannis Lambrou, Vassilis Vassiliades, and Chris Christodoulou</i>	

Apprenticeship and Inverse Reinforcement Learning

Bayesian Multitask Inverse Reinforcement Learning	273
<i>Christos Dimitrakakis and Constantin A. Rothkopf</i>	
Batch, Off-Policy and Model-Free Apprenticeship Learning	285
<i>Edouard Klein, Matthieu Geist, and Olivier Pietquin</i>	

Real-World Reinforcement Learning

Introduction of Fixed Mode States into Online Profit Sharing and Its Application to Waist Trajectory Generation of Biped Robot	297
<i>Seiya Kuroda, Kazuteru Miyazaki, and Hiroaki Kobayashi</i>	
MapReduce for Parallel Reinforcement Learning	309
<i>Yuxi Li and Dale Schuurmans</i>	
Compound Reinforcement Learning: Theory and an Application to Finance	321
<i>Tohgoroh Matsui, Takashi Goto, Kiyoshi Izumi, and Yu Chen</i>	
Proposal and Evaluation of the Active Course Classification Support System with Exploitation-Oriented Learning	333
<i>Kazuteru Miyazaki and Masaaki Ida</i>	

Author Index	345
--------------------	-----