Lecture Notes in Artificial Intelligence 9782

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

More information about this series at http://www.springer.com/series/1244

Bas Steunebrink · Pei Wang
Ben Goertzel (Eds.)

Artificial General Intelligence

9th International Conference, AGI 2016 New York, NY, USA, July 16–19, 2016 Proceedings

Springer

Editors
Bas Steunebrink
IDSIA
Manno
Switzerland

Ben Goertzel
Hong Kong Polytechnic University
Hong Kong
Hong Kong

Pei Wang
Temple University
Phoenixville, PA
USA

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Artificial Intelligence
ISBN 978-3-319-41648-9 ISBN 978-3-319-41649-6 (eBook)
DOI 10.1007/978-3-319-41649-6

Library of Congress Control Number: 2016943036

LNCS Sublibrary: SL7 – Artificial Intelligence

© Springer International Publishing Switzerland 2016
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, 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.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

This year marked the 60th anniversary of the "Dartmouth Summer Research Project on artificial intelligence" (1956), which launched artificial intelligence (AI) as a field of research. The original goal of AI was to replicate intelligence in machines; however, as the immense magnitude and difficulty of replicating human-level general intelligence soon became clear, AI fragmented into many sub-fields studying what we now call narrow-AI applications. Although the efforts of these sub-fields brought us extremely useful tools that now pervade virtually all technologies, efforts to work toward the original goal remained few and far between. In order to stimulate a return to the original goal of AI, a new name and corresponding conference series was created: Artificial General Intelligence (AGI). First organized in 2008, we are now in the ninth year of the AGI conference series.

To mark the 60th anniversary of AI as a field, the AGI 2016 conference was held as part of the larger HLAI 2016 event (the Joint Multi-Conference on Human-Level Intelligence), which co-located AGI 2016 with three other related conferences: BICA 2016 (the Annual International Conferences on Biologically Inspired Cognitive Architectures), NeSy 2016 (the Workshop Series on Neural-Symbolic Learning and Reasoning), and AIC 2016 (the Workshop Series on Artificial Intelligence and Cognition). Moreover, AGI 2016 was held back-to-back with IJCAI 2016 (the 25th International Joint Conference on Artificial Intelligence).

This volume contains the research papers presented at AGI 2016: The 9th Conference on Artificial General Intelligence, held during July 16–19, 2016 in New York City. In total, 67 papers were submitted to the conference. After each paper was reviewed by three Program Committee members, it was decided to accept 24 long papers and two short papers (39 % acceptance) for oral presentation, as well as ten papers for poster presentation.

In addition to these contributed talks, keynote speeches were shared with the larger HLAI event, and were delivered by Stephen Grossberg (Boston University), Gary Marcus (New York University and Geometric Intelligence Inc.), John Laird (University of Michigan), and David Aha (Naval Research Laboratory, Navy Center for Applied Research in Artificial Intelligence). Finally, the AGI 2016 conference featured two workshops, with the topics "Can Deep Neural Networks Solve the Problems of Artificial General Intelligence?" and "Environments and Evaluation for AGI."

July 2016

Bas Steunebrink
Pei Wang
Ben Goertzel

Organization

Organizing Committee

Ben Goertzel	OpenCog Foundation and Hanson Robotics, Hong Kong (Conference Chair)
Pei Wang	Temple University, USA
Bas Steunebrink	IDSIA, Switzerland
Matthew Iklé	Adams State University, USA
José Hernández-Orallo	Universitat Politècnica de València, Spain (Tutorials and Workshops Chair)
Brandon Rohrer	Microsoft (Poster and Demo Sessions Chair)
Ed Keller	New School (Local Co-chair)

Program Chairs

Pei Wang	Temple University, USA
Bas Steunebrink	IDSIA, Switzerland

Program Committee

Tsvi Achler	IBM Research Almaden, USA
Bo An	Nanyang Technological University, Singapore
Joscha Bach	MIT Media Lab, USA
Eric Baum	Baum Research Enterprises, USA
Tarek Richard Besold	Free University of Bozen-Bolzano, Italy
Jordi Bieger	Reykjavik University, Iceland
Dietmar Bruckner	Vienna University of Technology, Austria
Cristiano Castelfranchi	Institute of Cognitive Sciences and Technologies, Italy
Antonio Chella	Università di Palermo, Italy
Haris Dindo	Yewno Inc., USA
Benya Fallenstein	Machine Intelligence Research Institute, USA
Stan Franklin	University of Memphis, USA
Nil Geisweiller	Novamente LLC, USA
Ben Goertzel	OpenCog Foundation and Hanson Robotics, Hong Kong
Klaus Greff	IDSIA, Switzerland
José Hernández-Orallo	Universitat Politècnica de València, Spain
Bill Hibbard	University of Wisconsin - Madison, USA
Marcus Hutter	Australian National University, Australia
Matt Iklé	Adams State University, USA
Benjamin Johnston	University of Technology Sydney, Australia

Randal Koene	Carbon Copies, USA
Ramana Kumar	University of Cambridge, UK
Oliver Kutz	Free University of Bozen-Bolzano, Italy
Kai-Uwe Kühnberger	University of Osnabrück, Germany
Moshe Looks	Google Inc., USA
Maricarmen Martinez	Universidad de los Andes, Colombia
Amedeo Napoli	LORIA Nancy, France
Eric Nivel	Icelandic Institute for Intelligent Machines, Iceland
Sergei Obiedkov	National Research University Higher School of Economics, Russia
Laurent Orseau	Google DeepMind, UK
Guenther Palm	University of Ulm, Germany
Maxim Peterson	ITMO University, Russia
Florin Popescu	Fraunhofer FIRST, Germany
Alexey Potapov	AIDEUS, Russia
Paul S. Rosenbloom	University of Southern California, USA
Sebastian Rudolph	Technische Universität Dresden, Germany
Rafal Rzepka	Hokkaido University, Japan
Samer Schaat	Vienna University of Technology, Austria
Ute Schmid	University of Bamberg, Germany
Jürgen Schmidhuber	IDSIA, Switzerland
Zhongzhi Shi	Chinese Academy of Sciences, China
Leslie Smith	University of Stirling, UK
Javier Snaider	Google Inc., USA
Nate Soares	Machine Intelligence Research Institute, USA
Rupesh Srivastava	IDSIA, Switzerland
Bas Steunebrink	IDSIA, Switzerland
Claes Strannegård	University of Gothenburg, Sweden
Kristinn R. Thorisson	Reykjavik University, Iceland
Mario Verdicchio	Università degli Studi di Bergamo, Italy
Pei Wang	Temple University, USA
Roman Yampolskiy	University of Louisville, USA
Byoung-Tak Zhang	Seoul National University, South Korea

Steering Committee

Ben Goertzel	OpenCog Foundation and Hanson Robotics, Hong Kong (Chair)
Marcus Hutter	Australian National University, Australia

Contents

Self-Modification of Policy and Utility Function in Rational Agents 1
Tom Everitt, Daniel Filan, Mayank Daswani, and Marcus Hutter

Avoiding Wireheading with Value Reinforcement Learning 12
Tom Everitt and Marcus Hutter

Death and Suicide in Universal Artificial Intelligence 23
Jarryd Martin, Tom Everitt, and Marcus Hutter

Ultimate Intelligence Part II: Physical Complexity and Limits of Inductive Inference Systems 33
Eray Özkural

Open-Ended Intelligence: On the Role of Individuation in AGI. 43
David (Weaver) Weinbaum and Viktoras Veitas

The AGI Containment Problem. 53
James Babcock, János Kramár, and Roman Yampolskiy

Imitation Learning as Cause-Effect Reasoning 64
Garrett Katz, Di-Wei Huang, Rodolphe Gentili, and James Reggia

Some Theorems on Incremental Compression. 74
Arthur Franz

Rethinking Sigma's Graphical Architecture: An Extension to Neural Networks 84
Paul S. Rosenbloom, Abram Demski, and Volkan Ustun

Real-Time GA-Based Probabilistic Programming in Application to Robot Control 95
Alexey Potapov, Sergey Rodionov, and Vita Potapova

About Understanding. 106
Kristinn R. Thórisson, David Kremelberg, Bas R. Steunebrink, and Eric Nivel

Why Artificial Intelligence Needs a Task Theory: And What It Might Look Like 118
Kristinn R. Thórisson, Jordi Bieger, Thröstur Thorarensen, Jóna S. Sigurðardóttir, and Bas R. Steunebrink

Growing Recursive Self-Improvers 129
Bas R. Steunebrink, Kristinn R. Thórisson, and Jürgen Schmidhuber

Different Conceptions of Learning: Function Approximation vs. Self-Organization 140
Pei Wang and Xiang Li

The Emotional Mechanisms in NARS 150
Pei Wang, Max Talanov, and Patrick Hammer

The OpenNARS Implementation of the Non-Axiomatic Reasoning System 160
Patrick Hammer, Tony Lofthouse, and Pei Wang

Integrating Symbolic and Sub-symbolic Reasoning 171
Claes Strannegård and Abdul Rahim Nizamani

Integrating Axiomatic and Analogical Reasoning 181
Claes Strannegård, Abdul Rahim Nizamani, and Ulf Persson

Embracing Inference as Action: A Step Towards Human-Level Reasoning . . . 192
John Licato and Maxwell Fowler

Asymptotic Logical Uncertainty and the Benford Test 202
Scott Garrabrant, Tsvi Benson-Tilsen, Siddharth Bhaskar, Abram Demski, Joanna Garrabrant, George Koleszarik, and Evan Lloyd

Towards a Computational Framework for Function-Driven Concept Invention 212
Nico Potyka, Danny Gómez-Ramírez, and Kai-Uwe Kühnberger

System Induction Games and Cognitive Modeling as an AGI Methodology 223
Sean Markan

Integrating Model-Based Prediction and Facial Expressions in the Perception of Emotion 234
Nutchanon Yongsatianchot and Stacy Marsella

A Few Notes on Multiple Theories and Conceptual Jump Size 244
Grace Solomonoff

Generalized Temporal Induction with Temporal Concepts in a Non-axiomatic Reasoning System 254
Tony Lofthouse and Patrick Hammer

Introspective Agents: Confidence Measures for General Value Functions 258
Craig Sherstan, Adam White, Marlos C. Machado, and Patrick M. Pilarski

Automatic Sampler Discovery via Probabilistic Programming and Approximate Bayesian Computation 262
Yura Perov and Frank Wood

How Much Computation and Distributedness is Needed in Sequence Learning Tasks? 274
Mrwan Margem and Ozgur Yilmaz

Analysis of Algorithms and Partial Algorithms 284
Andrew MacFie

Estimating Cartesian Compression via Deep Learning 294
András Lőrincz, András Sárkány, Zoltán Á. Milacski, and Zoltán Tősér

A Methodology for the Assessment of AI Consciousness 305
Harry H. Porter III

Toward Human-Level Massively-Parallel Neural Networks with Hodgkin-Huxley Neurons 314
Lyle N. Long

Modeling Neuromodulation as a Framework to Integrate Uncertainty in General Cognitive Architectures 324
Frédéric Alexandre and Maxime Carrere

Controlling Combinatorial Explosion in Inference via Synergy with Nonlinear-Dynamical Attention Allocation 334
Ben Goertzel, Misgana Bayetta Belachew, Matthew Ikle', and Gino Yu

Probabilistic Growth and Mining of Combinations: A Unifying Meta-Algorithm for Practical General Intelligence 344
Ben Goertzel

Ideas for a Reinforcement Learning Algorithm that Learns Programs 354
Susumu Katayama

Author Index 363