

Genetic and Evolutionary Computation

Series Editors:

David E. Goldberg, ThreeJoy Associates, Inc., Urbana, IL, USA

John R. Koza, Stanford University, Los Altos, CA, USA

More information about this series at <http://www.springer.com/series/7373>

Wolfgang Banzhaf • Randal S. Olson
William Tozier • Rick Riolo
Editors

Genetic Programming Theory and Practice XV

Editors

Wolfgang Banzhaf
BEACON Center for the Study
of Evolution in Action and
Department of Computer Science
Michigan State University
East Lansing, MI, USA

William Tozier
Institute for Biomedical Informatics
University of Pennsylvania
Philadelphia, PA, USA

Randal S. Olson
Institute for Biomedical Informatics
University of Pennsylvania
Philadelphia, PA, USA

Rick Riolo
Center for the Study of Complex Systems
University of Michigan
Ann Arbor, MI, USA

ISSN 1932-0167

Genetic and Evolutionary Computation

ISBN 978-3-319-90511-2

ISBN 978-3-319-90512-9 (eBook)

<https://doi.org/10.1007/978-3-319-90512-9>

Library of Congress Control Number: 2018941244

© Springer International Publishing AG, part of Springer Nature 2018

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. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature.

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

It was a great pleasure giving our keynotes at GPTP 2017. We thank the organizers for inviting us. One of us (Jeff Clune) had always heard of and been interested in attending GPTP, and the other of us (Ken Stanley) had participated before and looked forward to returning. We both greatly enjoyed the workshop. The gathering is unique in its size and format, enabling longer, deeper conversations about the ideas presented than most conferences or workshops allow for. It was exciting to both see many senior leaders in the field, including many old friends, and meet the new researchers who have recently entered the field and are doing great work.

For one of us (Jeff), the setting of this workshop was personally meaningful. I attended the University of Michigan as an undergraduate and had my first taste of research working with Carl Simon, who is the founding director of the Center for the Study of Complex Systems, which is the official host of GPTP. Additionally, when looking for the best university at which to conduct my PhD research, I met with Rick Riolo, long-time GPTP organizer, in the very building GPTP was held in this year. It was great to return to Ann Arbor and that specific building so many years later, thinking back to my origins as a researcher and my desire to dedicate my career to exactly the questions we collectively discussed and study. We are all fortunate to be able to spend our lives researching such fascinating topics. It was especially nice to have Carl visit with the group and for me to be able to personally thank him for kick-starting my research career so many years ago. It is unfortunate Rick could not attend, but it was heartwarming to hear how much of an impact he has had on the GPTP community and how appreciated he is personally and professionally by all GPTP attendees.

Many of the ideas that were presented and discussed were exciting and innovative and have great potential. One that strongly resonated with both of us is Lexicase selection by Lee Spector and his students. One of us (Jeff), along with his PhD student Joost Huizinga, independently invented a similar idea in the combinatorial multi-objective evolutionary algorithm (CMOEA). A great function of conferences is learning about what others in the field are working on and discussing those ideas. Here was a perfect example. Prior to GPTP, Joost and I were not aware of Lexicase selection, but the format of the workshop both allowed us to learn about it during

Lee's presentation and afforded time for a prolonged discussion of it and how it relates to, and differs from, CMOEA. Joost and I have since experimented with Lexicase selection and have added it as a comparison algorithm to our upcoming paper on CMOEA, showing one immediate impact on research the workshop has already had. More generally, both Ken and I love the ideas behind Lexicase and CMOEA and believe they can propel the field toward producing more robust, generalist solutions that solve not just one problem, but many.

There were a multitude of interesting ideas presented and discussed, and we do not have space here to list them all. We will instead quickly mention a few more that particularly resonated with us. One is the work by Randy Olson and Jason Moore on AutoML. We think it is high time to finally realize the long-standing goal of automating machine learning pipelines. Doing so will both expand the impact of machine learning throughout society and catalyze faster progress in machine learning research. A second idea that resonated with us is the Eco-EA algorithm that Charles Ofria presented on. Jeff recalls the original version of that algorithm, which was developed by Sherri Goings, Charles, and collaborators at Michigan State while he was a PhD student. The idea tries to abstract one driver of natural diversity, which is having multiple niches, each with limited resources. If too many agents are currently exploiting one niche (e.g., solving one problem, or solving it in a particular way), there is a reward for some agents in the population to become different in order to exploit less depleted resources.

In fact, we believe that inventing open-ended evolution, in which a computational process endlessly creates an increasingly large set of diverse, high-quality solutions, is one of the great open scientific challenges. We also believe GPTP and similar evolutionary algorithm communities have the potential to make major advances toward this goal. A substantial amount of our own research, including much of what we both presented on in our keynotes, is dedicated to this quest. The Eco-EA algorithm has inspired some of this work, and we are delighted to see that it is continuing to be investigated and enhanced. We would be particularly excited to see how it can be improved to automatically allow niches to be created in a truly open-ended way. Furthermore, at GPTP, some of the best parts of the gathering include the side discussions, and open-endedness was a big topic in those conversations. For Ken, engaging with attendees about some of the hard questions in open-endedness genuinely broadened his appreciation for and understanding of the problem, even after studying it for years. That is the kind of outcome that an intimate and extended gathering like GPTP can offer that other more conventional venues rarely reproduce. We both hope that GPTP can become a catalyst for the growth of open-endedness as a field and community.

One of the discussions at the workshop was how to emulate the success of the community that trains deep neural networks via deep learning. One of us (Jeff) suggested then that what caused the world to take notice of, celebrate, and heavily invest in deep learning is simply that it works extremely well on hard problems. He quoted Steve Martin, who says "be so good they can't ignore you." One example of that was presented at GPTP by Michael Korn's. His hedge fund, Korn's Associates, built software for investing based on genetic programming, and Korn's credited

GPTP as crucial in the development of this software. Korn's Associates sold some rights to use this software to Lantern Credit for cash and shares valued at \$4.5 million USD. That is a great success story for the community and an example of building something that works so well that it cannot be ignored. Many of the ideas above, and the others presented at the workshop and described in these proceedings, also have the potential to deliver impressive, impossible-to-ignore results. Now the hard work begins to show their true potential. That requires hard science, which inevitably includes work on diagnostic (aka "toy") problems. However, and importantly, it also requires that we increasingly shoot for the stars. That means solving problems so challenging that the world will be forced to take notice of the wonderful work being done by this small but dedicated community. On that note, let's roll up our sleeves, set our ambitions high, and get to work!

Loy and Edith Harris Associate Professor in Computer Science

Jeff Clune

Director, Evolving Artificial Intelligence Lab

University of Wyoming

Senior Research Scientist, Uber AI Labs

San Francisco, CA, USA

Professor in Computer Science

Kenneth O. Stanley

Director, Evolutionary Complexity (EPlex) Lab

University of Central Florida

Senior Research Scientist, Uber AI Labs

San Francisco, CA, USA

February 2018

Preface

The book you hold in hand is the proceedings of the Fifteenth Workshop on Genetic Programming Theory and Practice, an invitation-only workshop held from May 18–20, 2017, at the University of Michigan in Ann Arbor, MI, under the auspices of the Center for the Study of Complex Systems. Since 2003, this annual workshop has been a forum for theorists in and practical users of genetic programming to exchange ideas, share insights, and trade observations. The workshop is intentionally organized as an event, where speculation is welcome and where participants are encouraged to discuss ideas or results that are not necessarily ready for publication in peer-reviewed publication, or have been published in different places and are summarized in the contributions provided for presentation here.

In addition to our regular sessions and interspersed with discussion sessions were three invited keynote talks. While regular talks are usually 40 min to present ideas and take questions, keynote talks are 60 min presentation time plus 10 min for an immediate question and answer session. Often, the ideas of keynote talks provide the start for more in-depth discussions during our discussion sessions.

This year, the first keynote speaker was Jeff Clune, University of Wyoming and Uber AI Lab, with “A talk in two parts: AI Neuroscience, and Harnessing Illumination Algorithms.” This talk presented what is at the forefront of what modern AI has to offer these days, mostly through the deep learning technology: very efficient pattern recognition algorithms, in neural network type representations, which are not directly conducive to forming an understanding of what is actually going on. Jeff and his collaborators have come up with a method to examine these deep neural networks to tease out what they have learned in a particular domain, so as to understand and be able to predict what would happen if other patterns were fed into those networks.

The second keynote talk was by Kenneth Stanley of the University of Central Florida and Uber AI on “New Directions in Open-Ended Evolution.” Ken, who presently heads the AI lab of Uber, discussed his work on novelty search, and how the benefits of moving away from a simple fitness goal could inform genetic programming and allow it to come up with more creative solutions to problems by promoting diversity through fostering behavioral novelty. His most recent work on

minimal criterion evolution featured prominently in his talk and provided plenty of fodder for discussions on open-ended evolution.

The third keynote talk was presented by Patrick Shafto from the Department of Mathematics and Computer Science at Rutgers University. His talk, entitled “Cooperative Inference in Humans and Machines,” addressed a very important new development in AI/ML—the collaboration of humans and computers to extract information and produce knowledge from data. It turns out that in the age of big data this cooperation is much more efficient in producing valuable insights than either computer algorithms or human learning.

We hope that the contributions published in this collection provide an exciting snapshot of what is going on in genetic programming!

Acknowledgements

We would like to thank all of the participants for again making GP Theory and Practice a successful workshop 2017. As is always the case, it produced a lot of interesting and high-energy discussions, as well as speculative thoughts and new ideas for further work. The keynote speakers did an excellent job at raising our awareness and provided thought-provoking ideas about the potential of genetic programming and its place in the world.

We would also like to thank our financial supporters for making the existence of GP Theory and Practice possible for the past 15 years, and counting. For 2017, these include:

- The Center for the Study of Complex Systems at the University of Michigan, and especially Carl Simon and Charles Doering, the champions of the workshop series
- John Koza
- Michael F. Korn, Lantern LLC
- Stuart W. Card
- Thomas Kern

A number of people made key contributions to the organization and assisted our participants during their stay in Ann Arbor. Foremost among them are Linda Wood and Mita Gibson who made the workshop run smoothly with their diligent efforts behind the scene before, during, and after the workshop. Special thanks go to the Springer Publishing Company, for providing the editorial assistance for producing this book. We are particularly grateful for contractual assistance by Melissa Fearon at Springer and all their staff has done to make this book possible.

East Lansing, MI, USA
 Philadelphia, PA, USA
 Ann Arbor, MI, USA
 Ann Arbor, MI, USA
 February 2018

Wolfgang Banzhaf
 Randall S. Olson
 William Tozier
 Rick Riolo

Contents

| | | |
|----------|---|------------|
| 1 | Exploiting Subprograms in Genetic Programming | 1 |
| | Steven B. Fine, Erik Hemberg, Krzysztof Krawiec, and Una-May O'Reilly | |
| 2 | Schema Analysis in Tree-Based Genetic Programming | 17 |
| | Bogdan Burlacu, Michael Affenzeller, Michael Kommenda, Gabriel Kronberger, and Stephan Winkler | |
| 3 | Genetic Programming Symbolic Classification: A Study | 39 |
| | Michael F. Korn | |
| 4 | Problem Driven Machine Learning by Co-evolving Genetic Programming Trees and Rules in a Learning Classifier System | 55 |
| | Ryan J. Urbanowicz, Ben Yang, and Jason H. Moore | |
| 5 | Applying Ecological Principles to Genetic Programming | 73 |
| | Emily Dolson, Wolfgang Banzhaf, and Charles Ofria | |
| 6 | Lexicase Selection with Weighted Shuffle | 89 |
| | Sarah Anne Troise and Thomas Helmuth | |
| 7 | Relaxations of Lexicase Parent Selection | 105 |
| | Lee Spector, William La Cava, Saul Shanabrook, Thomas Helmuth, and Edward Pantridge | |
| 8 | A System for Accessible Artificial Intelligence | 121 |
| | Randal S. Olson, Moshe Sipper, William La Cava, Sharon Tartarone, Steven Vitale, Weixuan Fu, Patryk Orzechowski, Ryan J. Urbanowicz, John H. Holmes, and Jason H. Moore | |
| 9 | Genetic Programming Based on Error Decomposition: A Big Data Approach | 135 |
| | Amirhessam Tahmassebi and Amir H. Gandomi | |

| | |
|--|------------|
| 10 One-Class Classification of Low Volume DoS Attacks with Genetic Programming..... | 149 |
| Stjepan Picek, Erik Hemberg, Domagoj Jakobovic, and Una- May O'Reilly | |
| 11 Evolution of Space-Partitioning Forest for Anomaly Detection | 169 |
| Zhiruo Zhao, Stuart W. Card, Kishan G. Mehrotra, and Chilukuri K. Mohan | |
| Index..... | 185 |

Contributors

Michael Affenzeller Heuristic and Evolutionary Algorithms Laboratory, University of Applied Sciences Upper Austria, Hagenberg, Austria

Institute for Formal Models and Verification, Johannes Kepler University, Linz, Austria

Wolfgang Banzhaf BEACON Center for the Study of Evolution in Action and Department of Computer Science, Michigan State University, East Lansing, MI, USA

Bogdan Burlacu Heuristic and Evolutionary Algorithms Laboratory, University of Applied Sciences Upper Austria, Hagenberg, Austria

Institute for Formal Models and Verification, Johannes Kepler University, Linz, Austria

Suart W. Card Syracuse University, Syracuse, NY, USA

Emily Dolson BEACON Center for the Study of Evolution in Action and Department of Computer Science and Ecology, Evolutionary Biology, and Behavior Program, Michigan State University, East Lansing, MI, USA

Steven B. Fine MIT CSAIL, Cambridge, MA, USA

Weixuan Fu Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, USA

Amir H. Gandomi School of Business, Stevens Institute of Technology, Hoboken, NJ, USA

BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, MI, USA

Thomas Helmuth Hamilton College, Clinton, NY, USA

Erik Hemberg MIT CSAIL, Cambridge, MA, USA

John H. Holmes Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, USA

Domagoj Jakobovic University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia

Michael Kommenda Heuristic and Evolutionary Algorithms Laboratory, University of Applied Sciences Upper Austria, Hagenberg, Austria

Institute for Formal Models and Verification, Johannes Kepler University, Linz, Austria

Michael F. Korn Lantern Credit LLC, Henderson, NV, USA

Krzysztof Krawiec Poznan Institute of Technology, Poznań, Poland

Gabriel Kronberger Heuristic and Evolutionary Algorithms Laboratory, University of Applied Sciences Upper Austria, Hagenberg, Austria

William La Cava Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, USA

Kishan G. Mehrotra Syracuse University, Syracuse, NY, USA

Chilukuri K. Mohan Syracuse University, Syracuse, NY, USA

Jason H. Moore Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, USA

Charles Ofria BEACON Center for the Study of Evolution in Action and Department of Computer Science and Ecology, Evolutionary Biology, and Behavior Program, Michigan State University, East Lansing, MI, USA

Randal S. Olson Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, USA

Una-May O'Reilly MIT CSAIL, Cambridge, MA, USA

Patryk Orzechowski Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, USA

Department of Automatics and Biomedical Engineering, AGH University of Science and Technology, Krakow, Poland

Edward Pantridge MassMutual, Amherst, MA, USA

Stjepan Picek MIT CSAIL, Cambridge, MA, USA

Saul Shanabrook University of Massachusetts, Amherst, MA, USA

Moshe Sipper Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, USA

Department of Computer Science, Ben-Gurion University, Beer-Sheva, Israel

Lee Spector Hampshire College, Amherst, MA, USA

Amirhessam Tahmassebi Department of Scientific Computing, Florida State University, Tallahassee, FL, USA

Sharon Tartarone Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, USA

Sarah Anne Troise Washington and Lee University, Lexington, VA, USA

Ryan J. Urbanowicz Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, USA

Steven Vitale Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, USA

Stephan Winkler Heuristic and Evolutionary Algorithms Laboratory, University of Applied Sciences Upper Austria, Hagenberg, Austria

Institute for Formal Models and Verification, Johannes Kepler University, Linz, Austria

Ben Yang Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, USA

Zhiruo Zhao Syracuse University, Syracuse, NY, USA