

Rick Riolo, Terence Soule and Bill Worzel (Eds.)

Genetic Programming Theory and Practice V

Genetic and Evolutionary Computation Series

Series Editors

David E. Goldberg
Consulting Editor
IlliGAL, Dept. of General Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61801 USA
Email: deg@uiuc.edu

John R. Koza
Consulting Editor
Medical Informatics
Stanford University
Stanford, CA 94305-5479 USA
Email: john@johnkoza.com

Selected titles from this series:

Markus Brameier, Wolfgang Banzhaf
Linear Genetic Programming, 2007
ISBN 978-0-387-31029-9

Nikolay Y. Nikolaev, Hitoshi Iba
Adaptive Learning of Polynomial Networks, 2006
ISBN 978-0-387-31239-2

Tetsuya Higuchi, Yong Liu, Xin Yao
Evolvable Hardware, 2006
ISBN 978-0-387-24386-3

David E. Goldberg
The Design of Innovation: Lessons from and for Competent Genetic Algorithms, 2002
ISBN 978-1-4020-7098-3

John R. Koza, Martin A. Keane, Matthew J. Streeter, William Mydlowec, Jessen Yu, Guido Lanza
Genetic Programming IV: Routine Human-Computer Machine Intelligence
ISBN: 978-1-4020-7446-2 (hardcover), 2003; ISBN: 978-0-387-25067-0 (softcover), 2005

Carlos A. Coello Coello, David A. Van Veldhuizen, Gary B. Lamont
Evolutionary Algorithms for Solving Multi-Objective Problems
ISBN: 978-0-306-46762-2, 2002; ISBN 978-0-387-33254-3 (2nd Edition), 2007

Lee Spector
Automatic Quantum Computer Programming: A Genetic Programming Approach
ISBN: 978-1-4020-7894-1 (hardcover), 2004; ISBN 978-0-387-36496-4 (softcover), 2007

William B. Langdon
Genetic Programming and Data Structures: Genetic Programming + Data Structures = Automatic Programming! 1998
ISBN: 978-0-7923-8135-8

For a complete listing of books in this series, go to <http://www.springer.com>

Rick Riolo
Terence Soule
Bill Worzel
(Eds.)

Genetic Programming Theory and Practice V

 Springer

Rick Riolo
Center for the Study of Complex Systems
323 West Hall
University of Michigan
Ann Arbor, MI 48109
rriolo@umich.edu

Terence Soule
Department of Computer Science
University of Idaho
Janssen Engineering Building
Moscow, ID 83844-1010
tsoule@cs.uidaho.edu

Bill Worzel
Genetics Squared
401 W. Morgan Rd.
Ann Arbor, MI 48108
billw@genetics2.com

Series Editors:

David E. Goldberg
Consulting Editor
IlliGAL, Dept. of General Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61801 USA
deg@uiuc.edu

John R. Koza
Consulting Editor
Medical Informatics
Stanford University
Stanford, CA 94305-5479 USA
john@johnkoza.com

Chapter 8 ©2008 Gregory S. Hornby. Used with permission.

Library of Congress Control Number: 2007937580

ISBN-13: 978-0-387-76307-1 e-ISBN-13: 978-0-387-76308-8

Printed on acid-free paper.

© 2008 Springer Science+Business Media, LLC

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

9 8 7 6 5 4 3 2 1

springer.com

Contents

Contributing Authors	vii
Preface	xi
Foreword	xiii
1	
Genetic Programming: Theory and Practice <i>Terence Soule, Rick L. Riolo and Bill Worzel</i>	1
2	
Better Solutions Faster: Soft Evolution of Robust Regression Models in Pareto Genetic Programming <i>Ekaterina Vladislavleva, Guido Smits and Mark Kotanchek</i>	13
3	
Manipulation of Convergence in Evolutionary Systems <i>Gearoid Murphy and Conor Ryan</i>	33
4	
Large-Scale, Time-Constrained Symbolic Regression-Classification <i>Michael F. Korns</i>	53
5	
Solving Complex Problems in Human Genetics Using Genetic Program- ming: The Importance of Theorist-Practitioner-Computer Interaction <i>Jason H. Moore, Nate Barney and Bill C. White</i>	69
6	
Towards an Information Theoretic Framework for Genetic Programming <i>Stuart W. Card and Chilukuri K. Mohan</i>	87
7	
Investigating Problem Hardness of Real Life Applications <i>Leonardo Vanneschi</i>	107
8	
Improving the Scalability of Generative Representations for Open-Ended Design <i>Gregory S. Hornby</i>	125

9

Program Structure-Fitness Disconnect and Its Impact On Evolution In Genetic Programming 143

A.A. Almal, C.D. MacLean and W.P. Worzel

10

Genetic Programming with Reuse of Known Designs for Industrially Scalable, Novel Circuit Design 159

Trent McConaghy, Pieter Palmers, Georges Gielen, and Michiel Steyaert

11

Robust Engineering Design of Electronic Circuits with Active Components Using Genetic Programming and Bond Graphs 185

Xiangdong Peng, Erik D. Goodman and Ronald C. Rosenberg

12

Trustable Symbolic Regression Models: Using Ensembles, Interval Arithmetic and Pareto Fronts to Develop Robust and Trust-Aware Models 201

Mark Kotanchek, Guido Smits and Ekaterina Vladislavleva

13

Improving Performance and Cooperation in Multi-Agent Systems 221

Terence Soule, Robert B. Heckendorn

14

An Empirical Study of Multi-Objective Algorithms for Stock Ranking 239

Ying L. Becker, Harold Fox, Peng Fei

15

Using GP and Cultural Algorithms to Simulate the Evolution of an Ancient Urban Center 261

Robert G. Reynolds, Mostafa Z. Ali and Patrick Franzel

Index

277

Contributing Authors

Arpit Arvindkumar Almal is an evolutionary engineer at Genetics Squared, Inc., a computational discovery company (aalmal@umich.edu).

Nate Barney is a Programmer in the Computational Genetics Laboratory at Dartmouth Medical School (nate.barney <nate.barney@Dartmouth.edu>).

Ying Becker is Vice President, Advanced Research Center, State Street Global Advisors, Boston, MA USA (Ying_Becker@ssga.com).

Stuart W. Card is a PhD candidate in the Dep't of Electrical Engineering and Comp. Sci. at Syracuse U. (cards@ntcnet.com).

Peng Fei is a quantitative research analyst for the advanced research center at State Street Global Advisors (SSgA), the investment management arm of State Street Corporation, Boston, MA (peng_fei@ssga.com).

Georges Gielen is a Professor of Electrical Engineering Katholieke Universiteit Leuven, Belgium (georges.gielen@esat.kuleuven.be).

Erik D. Goodman is Professor of Electrical and Computer Engineering and of Mechanical Engineering at Michigan State University (goodman@egr.msu.edu).

Gregory S. Hornby is a Project Scientist with the University of California Santa Cruz, working at NASA Ames Research Center in Moffett Field, CA (hornby@email.arc.nasa.gov).

Lee W. Jones is a researcher at Genetic Programming, Inc. in Mountain View, CA (lee@genetic-programming.com).

Michael F. Korn is Chief Technology Officer at Investment Science Corporation (mkorns@korns.com).

Mark E. Kotanchek is Chief Technology Officer of Evolved Analytics, a data modeling consulting and systems company (mark@evolved-analytics.com).

John R. Koza is Consulting Professor at Stanford University in the Department of Electrical Engineering (koza@stanford.edu).

Duncan MacLean is co-founder of Genetics Squared, Inc., a computational discovery company working in the pharmaceutical industry (dmaclean@acm.org).

Trent McConaghy is Chief Scientific Officer at Solido Design Automation Inc., as well as PhD candidate at ESAT-MICAS, Katholieke Universiteit Leuven, Belgium (trent.mcconaghy@yahoo.com).

Chilukuri K. Mohan is Professor of Electrical Engineering and Computer Science at Syracuse University (ckmohan@syr.edu).

Jason H. Moore is the Frank Lane Research Scholar in Computational Genetics and Associate Professor of Genetics at Dartmouth Medical School (Jason.H.Moore@Dartmouth.edu).

Gearoid Murphy is a Ph.D. student of Artificial Intelligence in the Biocomputing and Developmental systems research group of the University of Limerick, Ireland (gearoid.murphy@ul.ie).

Pieter Palmers is a PhD candidate at the Microelectronics and Sensors (ESAT-MICAS) Lab at Katholieke Universiteit Leuven, Belgium (pieter.palmers@esat.kuleuven.be).

Rick Riolo is Director of the Computer Lab and Associate Research Scientist in the Center for the Study of Complex Systems at the University of Michigan (rriolo@umich.edu).

Ronald C. Rosenberg is Professor of Mechanical Engineering at Michigan State University (roserber@egr.msu.edu).

Conor Ryan is Senior Lecturer in the Department of Computer Science and Information Systems at University of Limerick, Ireland where he leads the Biocomputing and Developmental Systems Group (conor.ryan@ul.ie).

Guido F. Smits is a Research and Development Leader in the Modelling Group within the Engineering and Process Sciences R&D Organization of the Dow Chemical Company (gfsmits@dow.com).

Terence Soule is an Associate Professor of Computer Science at the University of Idaho, Moscow, ID (tsoule@cs.uidaho.edu).

Michiel Steyaert is a Professor of Electrical Engineering Katholieke Universiteit Leuven, Belgium (michiel.steyaert@esat.kuleuven.be).

Leonardo Vanneschi is Assistant Professor at the University of Milano-Bicocca (Italy), in the Department of Informatics, Systems and Communication (D.I.S.Co.) (vanneschi@disco.unimib.it).

Ekaterina Vladislavleva is a Ph.D candidate in the Department of Operations Research of Tilburg University and in the R&D Organization of the Dow Benelux B.V., Netherlands (CVladislavleva@dow.com).

Bill C. White is a Senior Programmer in the Computational Genetics Laboratory at Dartmouth Medical School (Bill C. White <Bill.C.White@Dartmouth.edu>).

Bill Worzel is the Chief Technology Officer and co-founder of Genetics Squared, Inc., a computational discovery company working in the pharmaceutical industry (billw@genetics2.com).

Preface

The work described in this book was first presented at the Fifth Workshop on Genetic Programming, Theory and Practice, organized by the Center for the Study of Complex Systems at the University of Michigan, Ann Arbor, 17-19 May 2007. The goal of this workshop series is to promote the exchange of research results and ideas between those who focus on Genetic Programming (GP) theory and those who focus on the application of GP to various real-world problems. In order to facilitate these interactions, the number of talks and participants was small and the time for discussion was large. Further, participants were asked to review each other's chapters *before* the workshop. Those reviewer comments, as well as discussion at the workshop, are reflected in the chapters presented in this book. Additional information about the workshop, addendums to chapters, and a site for continuing discussions by participants and by others can be found at <http://cscs.umich.edu/gptp2007>.

We thank all the workshop participants for making the workshop an exciting and productive three days. In particular we thank all the authors, without whose hard work and creative talents, neither the workshop nor the book would be possible. We also thank our keynote speakers PZ Meyers, Associate Professor of Biology at the University of Minnesota, Morris, and Wolfgang Banzhaf, Chair and Professor of Computer Science at the Department of Computer Science of Memorial University of Newfoundland, Canada. Both keynotes delivered thought-provoking talks comparing and contrasting natural biological systems to Genetic Programming and to Artificial Evolution in general, all of which inspired a great deal of discussion among the participants.

The workshop received support from these sources:

- The Center for the Study of Complex Systems (CSCS);
- Third Millennium Venture Capital Limited;
- Michael Korn, Investment Science Corporation.
- State Street Global Advisors, Boston, MA;

- Biocomputing and Developmental Systems Group, Computer Science and Information Systems, University of Limerick;
- Christopher T. May, RedQueen Capital Management;
- Dow Chemical Corporation, Midland, MI; and
- Genetics Squared, Inc, Ann Arbor, MI.

We thank all of our sponsors for their kind and generous support for the workshop and GP research in general.

A number of people made key contributions to running the workshop and assisting the attendees while they were in Ann Arbor. Foremost among them was Howard Oishi, assisted by Sarah Cherng and Mike Bommarito. After the workshop, many people provided invaluable assistance in producing this book. Special thanks go to Sarah Cherng, who did a wonderful job working with the authors, editors and publishers to get the book completed very quickly. Thanks to William Tozier for assisting in copy-editing some of the chapters. Valerie Schofield and Melissa Fearon's editorial efforts were invaluable from the initial plans for the book through its final publication. Thanks also to Deborah Doherty of Springer for helping with various technical publishing issues. Finally, we thank Carl Simon, Director of CSCS, for his support for this endeavor from its very inception.

RICK RIOLO, TERENCE SOULE AND BILL WORZEL

Foreword

It was a great joy for me to be invited to the 5th Genetic Programming Workshop on Theory and Practice, held in May 2007 in Ann Arbor. The organizers are to be congratulated to a well conceived event. The Center for the Study of Complex Systems (CSCS) at the University of Michigan was a fabulous host again this year. Much as in earlier years, participants at the workshop are a unique blend of theoreticians and practitioners in GP, and the workshop is an ideal place to move forward with ideas as both streams fertilize each other.

Although I did participate at this workshop in earlier years, I was honoured this year to give a keynote speech, along with Developmental Biologist, Dr. PZ Myers, from the University of Minnesota, Morris campus.

I was particularly impressed this year by the dedication to community progress. Frequently it was said that ideas introduced and examined in particular algorithmic variants could and should be included in all tools for GP. Among those particularly mentioned this year were information-theoretic measures of fitness and diversity in a populations, as well as recipes like restart of runs and random sampling of fitness cases.

As a result of this community effort, the GPTP workshop has developed into a driving force for progress in GP. Some participants were already planning their prospective next year's contribution, with progress anticipated from integration of new ideas discussed at this year's workshop. If such momentum can be maintained, surely GPTP workshops will go down as some of the most influential events in the history of Genetic Programming.

The keynote speakers this year were selected with an eye on the impact of biological discoveries in Evolutionary Computation. Genetic Programming, as much as this seems strange, happens to be the closest algorithmic incorporation of natural evolution. It offers a rich set of by-products in its behavior, which baffle practitioners, yet are to be expected if one compares the algorithms to their natural counterparts.

A lot could be adopted from what Biology has learned over the past 20 years, and from the feedback I received on my talk there seems to be a general recognition that learning from Biology will infuse further innovations into our field and propel Genetic Programming forward in the coming years.

In my opinion, our algorithmic implementations have just started to explore the power of evolution (and development), and there is much more to be found in the coming years. I went away from GPTP reinvigorated and wish the reader the same from studying the contributions in this collection of talks.

Wolfgang Banzhaf, Head
Department of Computer Science
Memorial University of Newfoundland
St. John's, Canada
July, 2007