

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Alfred Kobsa

University of California, Irvine, CA, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

TU Dortmund University, Germany

Madhu Sudan

Microsoft Research, Cambridge, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max Planck Institute for Informatics, Saarbruecken, Germany

Mario Giacobini Leonardo Vanneschi
William S. Bush (Eds.)

Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics

10th European Conference, EvoBIO 2012
Málaga, Spain, April 11-13, 2012
Proceedings

Volume Editors

Mario Giacobini

University of Torino

Department of Animal Production Epidemiology and Ecology

Via Leonardo da Vinci 44, 10095 Grugliasco (TO), Italy

E-mail: mario.giacobini@unito.it

Leonardo Vanneschi

Universidade Nova de Lisboa, ISEGI

1070-312 Lisboa, Portugal

and

University of Milano-Bicocca, D.I.S.Co.

Viale Sarca 336, 20126 Milan, Italy

E-mail: lvanneschi@isegi.unl.pt

William S. Bush

Vanderbilt University, Center for Human Genetics Research

Department of Biomedical Informatics

519 Light Hall, Nashville, TN 37232, USA

E-mail: william.s.bush@vanderbilt.edu

Cover illustration:

"Chair No. 17" by The Painting Fool (www.thepaintingfool.com)

ISSN 0302-9743

e-ISSN 1611-3349

ISBN 978-3-642-29065-7

e-ISBN 978-3-642-29066-4

DOI 10.1007/978-3-642-29066-4

Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012933967

CR Subject Classification (1998): J.3, H.2.8, E.1, I.2, F.1, F.2.1

LNCS Sublibrary: SL 1 – Theoretical Computer Science and General Issues

© 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

Computational biology is a wide and varied discipline, incorporating aspects of statistical analysis, data structure and algorithm design, machine learning, and mathematical modeling toward the processing and improved understanding of biological data. Experimentalists now routinely generate new information on such a massive scale that the techniques of computer science are needed to establish any meaningful result. As a consequence, biologists now face the challenges of algorithmic complexity and tractability, and combinatorial explosion when conducting even basic analyses. The goal of the 10th European Conference on Evolutionary Computation, Machine Learning, and Data Mining in Computational Biology (EvoBIO 2012) was to bring together experts across multiple fields to discuss novel methods for tackling complex biological problems, and the beauty of EvoBIO is that often these experts draw inspiration from biological systems in order to produce solutions to biological problems.

The 10th EvoBIO conference was held in Málaga, Spain, during April 11–13, 2012, at the Computer Science School of the University of Málaga, Spain. EvoBIO 2012 was held jointly with the 15th European Conference on Genetic Programming (EuroGP 2012), the 12th European Conference on Evolutionary Computation in Combinatorial Optimization (EvoCOP 2012), the First International Conference and 10th European Event on Evolutionary and Biologically Inspired Music, Sound, Art, and Design (EvoMUSART 2012), and the European Conference on the Applications of Evolutionary Computation. Collectively, the conferences are organized under the name Evo* 2012 (www.evostar.org). EvoBIO, held annually as a workshop since 2003, became a conference in 2007 and it is now the premier European event for those interested in the interface between evolutionary computation, machine learning, and data mining in computational biology. All papers in these proceedings were presented at EvoBIO 2012 in oral or poster presentations, and were received in response to a call for papers soliciting a wide range of topics in the realm of biological data analysis and computational biology.

EvoBIO 2012 added a new submission format: four-page abstracts reviewing, discussing, or extending work previously published in a journal. This form was requested in addition to our existing submission formats: 12-page full research articles describing new developments in methodology, approach, and/or application, eight-page system demonstrations of active or developing systems in research or practice domains, and eight-page short reports describing new methodologies, approaches, and/or applications. After peer review, EvoBIO accepted 15 papers for oral presentation and 8 for poster presentation.

With the goal of sharing inspiration in mind, EvoBIO and EuroGP held a special joint session to celebrate the tenth anniversary of EvoBIO. In this session, advances in the field of genetic programming were applied to problems of

computational biology, and likewise, the unique mechanisms present in biological systems were used to create new extensions to the paradigm of genetic programming. This invigorating session created new collaborations and encouraged the development of new approaches and their application to the biological problem domain.

First and foremost, we thank all the authors who spent time and effort to generate the fantastic contributions to this body of work. We thank the members of the Program Committee for their expert evaluation and review of the submitted papers. We also thank many members of the Evo* community who worked tirelessly to ensure a smooth and successful conference event; Jennifer Willies from Edinburgh Napier University for her unwavering dedication as event coordinator and the Institute for Informatics and Digital Innovation at Edinburgh Napier University; and Penousal Machado from the University of Coimbra for his fantastic work as Publicity Chair. We owe special thanks to Carlos Cotta from the University of Málaga for his outstanding planning as local organizer and for bringing Evo* to the beautiful city of Málaga. We extend our gratitude to the School of Computer Science directed by Jose M. Troya and the School of Telecommunications directed by Antonio Puerta from the University of Málaga for hosting our event, and to the Málaga Convention Bureau for their support of this conference. We also deeply appreciate the work of Marc Schoenauer from INRIA in France and the MyReview team for providing the publication management system and technical support.

We hope you enjoy the fascinating research articles included in this volume, and we invite you to contribute your work to EvoBIO 2013.

April 2012

Mario Giacobini
Leonardo Vanneschi
William S. Bush

Organization

EvoBIO 2012, together with EuroGP 2012, EvoCOP 2012, EvoAPPLICATIONS 2012, and EvoMUSART 2012 was part of EVO* 2012, Europe's premier co-located events in the field of evolutionary computing.

Program Chairs

Mario Giacobini	University of Torino, Italy
Leonardo Vanneschi	Universidade Nova de Lisboa, Portugal
	University of Milano-Bicocca, Milan, Italy
William S. Bush	Vanderbilt University in Nashville, TN, USA

Local Chair

Carlos Cotta	University of Málaga, Italy
--------------	-----------------------------

Publicity Chair

Penousal Machado	University of Coimbra, Portugal
------------------	---------------------------------

Proceedings Chair

Mario Giacobini	University of Torino, Italy
-----------------	-----------------------------

Steering Committee

Elena Marchiori	Radboud University, Nijmegen, The Netherlands
Jason H. Moore	Dartmouth Medical School in Lebanon, NH, USA
Clara Pizzuti	ICAR-CNR, Italy
Marylyn Ritchie	Vanderbilt University, USA

Program Committee

Jesus S. Aguilar-Ruiz	Universidad Pablo de Olavide, Spain
Wolfgang Banzhaf	Memorial University of Newfoundland, Canada
Jacek Blazewicz	Poznan University of Technology, Poland

Erik Boczko	Vanderbilt University, USA
Ernesto Costa	University of Coimbra, Portugal
Federico Divina	Pablo de Olavide University Seville, Spain
Jitesh Dundas	Edencore Technologies, USA
Alex Freitas	University of Kent, UK
Raffaele Giancarlo	University of Palermo, Italy
Raul Giraldez Rojo	Pablo de Olavide University, Spain
Rosalba Giugno	University of Catania, Italy
Jin-Kao Hao	University of Angers, France
Tom Heskes	Radboud University Nijmegen, The Netherlands
Ting Hu	Dartmouth College, Hanover, USA
Zhenyu Jia	University of California, Irvine, USA
Mehmet Koyuturk	Case Western Reserve University, USA
Michael Lones	University of York, UK
Penousal Machado	University of Coimbra, Portugal
Bob MacCallum	Imperial College London, UK
Elena Marchiori	Radboud University, Nijmegen, The Netherlands
Andrew Martin	University College London, UK
Brett McKinney	University of Tulsa, USA
Jason H. Moore	Dartmouth College, Hanover, USA
Pablo Moscato	The University of Newcastle, Australia
Alison Motsinger-Reif	North Carolina State University, USA
Vincent Moulton	University of East Anglia, UK
Carlotta Orsenigo	Politecnico di Milano, Italy
Clara Pizzuti	ICAR-CNR, Italy
Paolo Provero	University of Torino, Italy
Michael Raymer	Wright State University, USA
Marylyn Ritchie	The Pennsylvania State University, USA
Simona Rombo	ICAR-CNR, Italy
Marc Schoenauer	LRI- Université Paris-Sud, France
Ugur Sezerman	Sabanci University, Turkey
Sara Silva	INESC-ID Lisboa, Portugal
Marc L. Smith	Vassar College, USA
El-Ghazali Talbi	Université des Sciences et Technologies de Lille, France
Marco Tomassini	University of Lausanne, Switzerland
Stephen Turner	University of Virginia, USA
Alfonso Urso	ICAR-CNR, Italy
Antoine van Kampen	Universiteit van Amsterdam, The Netherlands
Andreas Zell	University of Tübingen, Germany
Zhongming Zhao	Vanderbilt University, USA

Sponsoring Institutions

- School of Computer Science, University of Málaga, Spain
- School of Telecommunications, University of Málaga, Spain
- Málaga Convention Bureau, Spain
- The Institute for Informatics and Digital Innovationg, Edinburgh Napier University, UK

Table of Contents

Oral Contributions

Automatic Task Decomposition for the NeuroEvolution of Augmenting Topologies (NEAT) Algorithm	1
<i>Timmy Manning and Paul Walsh</i>	
Evolutionary Reaction Systems	13
<i>Luca Manzoni, Mauro Castelli, and Leonardo Vanneschi</i>	
Optimizing the Edge Weights in Optimal Assignment Methods for Virtual Screening with Particle Swarm Optimization	26
<i>Lars Rosenbaum, Andreas Jahn, and Andreas Zell</i>	
Lévy-Flight Genetic Programming: Towards a New Mutation Paradigm	38
<i>Christian Darabos, Mario Giacobini, Ting Hu, and Jason H. Moore</i>	
Understanding Zooplankton Long Term Variability through Genetic Programming	50
<i>Simone Marini and Alessandra Conversi</i>	
Inferring Disease-Related Metabolite Dependencies with a Bayesian Optimization Algorithm	62
<i>Holger Franken, Alexander Seitz, Rainer Lehmann, Hans-Ulrich Häring, Norbert Stefan, and Andreas Zell</i>	
A GPU-Based Multi-swarm PSO Method for Parameter Estimation in Stochastic Biological Systems Exploiting Discrete-Time Target Series . . .	74
<i>Marco S. Nobile, Daniela Besozzi, Paolo Cazzaniga, Giancarlo Mauri, and Dario Pescini</i>	
Tracking the Evolution of Cooperation in Complex Networked Populations	86
<i>Flávio L. Pinheiro, Francisco C. Santos, and Jorge M. Pacheco</i>	
GeNet: A Graph-Based Genetic Programming Framework for the Reverse Engineering of Gene Regulatory Networks	97
<i>Leonardo Vanneschi, Matteo Mondini, Martino Bertoni, Alberto Ronchi, and Mattia Stefano</i>	
Comparing Multiobjective Artificial Bee Colony Adaptations for Discovering DNA Motifs	110
<i>David L. González-Álvarez, Miguel A. Vega-Rodríguez, Juan A. Gómez-Pulido, and Juan M. Sánchez-Pérez</i>	

The Role of Mutations in Whole Genome Duplication	122
<i>Qinxin Pan, Christian Darabos, and Jason H. Moore</i>	
Comparison of Methods for Meta-dimensional Data Analysis Using in Silico and Biological Data Sets	134
<i>Emily R. Holzinger, Scott M. Dudek, Alex T. Frase, Brooke Fridley, Prabhakar Chalise, and Marylyn D. Ritchie</i>	
Inferring Phylogenetic Trees Using a Multiobjective Artificial Bee Colony Algorithm	144
<i>Sergio Santander-Jiménez, Miguel A. Vega-Rodríguez, Juan A. Gómez-Pulido, and Juan M. Sánchez-Pérez</i>	
Prediction of Mitochondrial Matrix Protein Structures Based on Feature Selection and Fragment Assembly	156
<i>Gualberto Asencio-Cortés, Jesús S. Aguilar-Ruiz, Alfonso E. Márquez-Chamorro, Roberto Ruiz, and Cosme E. Santiesteban-Toca</i>	
Poster Contributions	
Feature Selection for Lung Cancer Detection Using SVM Based Recursive Feature Elimination Method	168
<i>Kesav Kancherla and Srinivas Mukkamala</i>	
Measuring Gene Expression Noise in Early <i>Drosophila</i> Embryos: The Highly Dynamic Compartmentalized Micro-environment of the Blastoderm Is One of the Main Sources of Noise	177
<i>Alexander V. Spirov, Nina E. Golyandina, David M. Holloway, Theodore Alexandrov, Ekaterina N. Spirova, and Francisco J.P. Lopes</i>	
Artificial Immune Systems Perform Valuable Work When Detecting Epistasis in Human Genetic Datasets	189
<i>Delaney Granizo-Mackenzie and Jason H. Moore</i>	
A Biologically Informed Method for Detecting Associations with Rare Variants	201
<i>Carrie C. Buchanan, John R. Wallace, Alex T. Frase, Eric S. Torstenson, Sarah A. Pendergrass, and Marylyn D. Ritchie</i>	
Complex Detection in Protein-Protein Interaction Networks: A Compact Overview for Researchers and Practitioners	211
<i>Clara Pizzuti, Simona E. Rombo, and Elena Marchiori</i>	
Short-Range Interactions and Decision Tree-Based Protein Contact Map Predictor	224
<i>Cosme E. Santiesteban-Toca, Gualberto Asencio-Cortés, Alfonso E. Márquez-Chamorro, and Jesús S. Aguilar-Ruiz</i>	

A NSGA-II Algorithm for the Residue-Residue Contact Prediction	234
<i>Alfonso E. Márquez-Chamorro, Federico Divina,</i>	
<i>Jesús S. Aguilar-Ruiz, Jaume Bacardit,</i>	
<i>Gualberto Asencio-Cortés, and Cosme E. Santiesteban-Toca</i>	

Abstract Contributions

<i>In Silico</i> Infection of the Human Genome	245
<i>W.B. Langdon and M.J. Arno</i>	
Improving Phylogenetic Tree Interpretability by Means of Evolutionary Algorithms	250
<i>Francesco Cerutti, Luigi Bertolotti, Tony L. Goldberg, and</i>	
<i>Mario Giacobini</i>	
Author Index	255