

Advances in Intelligent and Soft Computing

Editor-in-Chief

Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

Further volumes of this series can be found on our homepage: springer.com

Vol. 58. J. Mehnen, A. Tiwari,
M. Köppen, A. Saad (Eds.)
Applications of Soft Computing, 2009
ISBN 978-3-540-89618-0

Vol. 59. K.A. Cyran,
S. Kozielski, J.F. Peters,
U. Stańczyk, A. Wakulicz-Deja (Eds.)
Man-Machine Interactions, 2009
ISBN 978-3-642-00562-6

Vol. 60. Z.S. Hippe,
J.L. Kulikowski (Eds.)
Human-Computer Systems Interaction, 2009
ISBN 978-3-642-03201-1

Vol. 61. W. Yu, E.N. Sanchez (Eds.)
*Advances in Computational
Intelligence, 2009*
ISBN 978-3-642-03155-7

Vol. 62. B. Cao,
T.-F. Li, C.-Y. Zhang (Eds.)
*Fuzzy Information and
Engineering Volume 2, 2009*
ISBN 978-3-642-03663-7

Vol. 63. Á. Herrero, P. Gastaldo,
R. Zunino, E. Corchado (Eds.)
*Computational Intelligence in Security for
Information Systems, 2009*
ISBN 978-3-642-04090-0

Vol. 64. E. Tkacz, A. Kapczynski (Eds.)
*Internet – Technical Development and
Applications, 2009*
ISBN 978-3-642-05018-3

Vol. 65. E. Kaçki, M. Rudnicki,
J. Stempczyńska (Eds.)
Computers in Medical Activity, 2009
ISBN 978-3-642-04461-8

Vol. 66. G.Q. Huang,
K.L. Mak, P.G. Maropoulos (Eds.)
*Proceedings of the 6th CIRP-Sponsored
International Conference on Digital
Enterprise Technology, 2009*
ISBN 978-3-642-10429-9

Vol. 67. V. Snášel, P.S. Szczepaniak,
A. Abraham, J. Kacprzyk (Eds.)
Advances in Intelligent Web Mastering - 2, 2010
ISBN 978-3-642-10686-6

Vol. 68. V.-N. Huynh, Y. Nakamori,
J. Lawry, M. Inuiguchi (Eds.)
*Integrated Uncertainty Management and
Applications, 2010*
ISBN 978-3-642-11959-0

Vol. 69. E. Pietka and J. Kawa (Eds.)
Information Technologies in Biomedicine, 2010
ISBN 978-3-642-13104-2

Vol. 70. XXX

Vol. 71. XXX

Vol. 72. J.C. Augusto, J.M. Corchado,
P. Novais, C. Analide (Eds.)
Ambient Intelligence and Future Trends, 2010
ISBN 978-3-642-13267-4

Vol. 73. J.M. Corchado, P. Novais,
C. Analide, J. Sedano (Eds.)
*Soft Computing Models in Industrial and
Environmental Applications, 5th International
Workshop (SOCO 2010), 2010*
ISBN 978-3-642-13160-8

Vol. 74. M.P. Rocha, F.F. Riverola, H. Shatkay,
J.M. Corchado (Eds.)
Advances in Bioinformatics
ISBN 978-3-642-13213-1

Miguel P. Rocha,
Florentino Fernández Riverola, Hagit Shatkay,
and Juan Manuel Corchado (Eds.)

Advances in Bioinformatics

4th International Workshop on Practical
Applications of Computational Biology
and Bioinformatics 2010 (IWPACBB 2010)

 Springer

Editors

Miguel P. Rocha
Dep. Informática / CCTC
Universidade do Minho
Campus de Gualtar
4710-057 Braga
Portugal

Hagit Shatkay
Computational Biology and
Machine Learning Lab
School of Computing
Queen's University Kingston
Ontario K7L 3N6
Canada
E-mail: shatkay@cs.queensu.ca

Florentino Fernández-Riverola
Escuela Superior de
Ingeniería Informática
Edificio Politécnico,
Despacho 408
Campus Universitario
As Lagoas s/n
32004 Ourense
Spain
E-mail: riverola@ei.uvigo.es

Juan Manuel Corchado
Departamento de Informática
y Automática
Facultad de Ciencias
Universidad de Salamanca
Plaza de la Merced S/N
37008 Salamanca
Spain
E-mail: corchado@usal.es

ISBN 978-3-642-13213-1

e-ISBN 978-3-642-13214-8

DOI 10.1007/978-3-642-13214-8

Advances in Intelligent and Soft Computing

ISSN 1867-5662

Library of Congress Control Number: Applied For

© 2010 Springer-Verlag Berlin Heidelberg

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, reuse of illustrations, recitation, broadcasting, reproduction on microfilm 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 for 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.

Typeset & Cover Design: Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper

5 4 3 2 1 0

springer.com

Preface

The fields of Bioinformatics and Computational Biology have been growing steadily over the last few years boosted by an increasing need for computational techniques that can efficiently handle the huge amounts of data produced by the new experimental techniques in Biology. This calls for new algorithms and approaches from fields such as Data Integration, Statistics, Data Mining, Machine Learning, Optimization, Computer Science and Artificial Intelligence.

Also, new global approaches, such as Systems Biology, have been emerging replacing the reductionist view that dominated biological research in the last decades. Indeed, Biology is more and more a science of information needing tools from the information technology field. The interaction of researchers from different scientific fields is, more than ever, of foremost importance and we hope this event will contribute to this effort.

IWPACBB'10 technical program included a total of 30 papers (26 long papers and 4 short papers) spanning many different sub-fields in Bioinformatics and Computational Biology. Therefore, the technical program of the conference will certainly be diverse, challenging and will promote the interaction among computer scientists, mathematicians, biologists and other researchers.

We would like to thank all the contributing authors, as well as the members of the Program Committee and the Organizing Committee for their hard and highly valuable work. Their work has helped to contribute to the success of the IWPACBB'10 event. IWPACBB'10 wouldn't exist without your contribution.

Miguel Rocha
Florentino Fdez-Riverola
IWPACBB'10 Organizing Co-chairs

Juan Manuel Corchado
Hagit Shatkay
IWPACBB'10 Programme Co-chairs

Organization

General Co-chairs

Miguel Rocha	University of Minho (Portugal)
Florentino Riverola	University of Vigo (Spain)
Juan M. Corchado	University of Salamanca (Spain)
Hagit Shatkay	Queens University, Ontario (Canada)

Program Committee

Juan M. Corchado (Co-chairman)	University of Salamanca (Spain)
Alicia Troncoso	Universidad of Pablo de Olavide (Spain)
Alípio Jorge	LIAAD/INESC, Porto LA (Portugal)
Anália Lourenço	University of Minho (Portugal)
Arlindo Oliveira	INESC-ID, Lisboa (Portugal)
Arlo Randall	University of California Irvine (USA)
B. Cristina Pelayo	University of Oviedo (Spain)
Christopher Henry	Argonne National Labs (USA)
Daniel Gayo	University of Oviedo (Spain)
David Posada	Univ. Vigo (Spain)
Emilio S. Corchado	University of Burgos (Spain)
Eugénio C. Ferreira	IBB/CEB, University of Minho (Portugal)
Fernando Diaz-Gómez	University of Valladolid (Spain)
Gonzalo Gómez-López	UBio/CNIO, Spanish National Cancer Research Centre (Spain)
Isabel C. Rocha	IBB/CEB, University of Minho (Portugal)
Jesús M. Hernández	University of Salamanca (Spain)
Jorge Vieira	IBMC, Porto (Portugal)
José Adserias	University of Salamanca (Spain)
José L. López	University of Salamanca (Spain)
José Luís Oliveira	Univ. Aveiro (Portugal)
Juan M. Cueva	University of Oviedo (Spain)
Júlio R. Banga	IIM/CSIC, Vigo (Spain)

Kaustubh Raosaheb Patil	Max-Planck Institute for Informatics(Germany)
Kiran R. Patil	Biocentrum, DTU (Denmark)
Lourdes Borrajo	University of Vigo (Spain)
Luis M. Rocha	Indiana University (USA)
Manuel J. Maña López	University of Huelva (Spain)
Margarida Casal	University of Minho (Portugal)
Maria J. Ramos	FCUP, University of Porto (Portugal)
Martin Krallinger	CNB, Madrid (Spain)
Nicholas Luscombe	EBI (UK)
Nuno Fonseca	CRACS/INESC, Porto (Portugal)
Oscar Sanjuan	University of Oviedo (Spain)
Paulo Azevedo	University of Minho (Portugal)
Paulino Gómez-Puertas	University Autónoma de Madrid (Spain)
Pierre Balde	University of California Irvine (USA)
Rui Camacho	LIACC/FEUP, University of Porto (Portugal)
Rui Brito	University of Coimbra (Portugal)
Rui C. Mendes	CCTC, University of Minho (Portugal)
Sara Madeira	IST/INESC, Lisboa (Portugal)
Sérgio Deusdado	IP Bragança (Portugal)
Vítor Costa	University of Porto (Portugal)

Organizing Committee

Miguel Rocha (Co-chairman)	CCTC, Univ. Minho (Portugal)
Florentino Fernández Riverola (Co-chairman)	University of Vigo (Spain)
Juan F. De Paz	University of Salamanca (Spain)
Daniel Glez-Peña	University of Vigo (Spain)
José P. Pinto	University of Minho (Portugal)
Rafael Carreira	University of Minho (Portugal)
Simão Soares	University of Minho (Portugal)
Paulo Vilaça	University of Minho (Portugal)
Hugo Costa	University of Minho (Portugal)
Paulo Maia	University of Minho (Portugal)
Pedro Evangelista	University of Minho (Portugal)
Óscar Dias	University of Minho (Portugal)

Contents

Microarrays

Highlighting Differential Gene Expression between Two Condition Microarrays through Heterogeneous Genomic Data: Application to <i>Leishmania infantum</i> Stages Comparison	1
<i>Liliana López Kleine, Víctor Andrés Vera Ruiz</i>	

An Experimental Evaluation of a Novel Stochastic Method for Iterative Class Discovery on Real Microarray Datasets ...	9
<i>Héctor Gómez, Daniel Glez-Peña, Miguel Reboiro-Jato, Reyes Pavón, Fernando Díaz, Florentino Fdez-Riverola</i>	

Automatic Workflow during the Reuse Phase of a CBP System Applied to Microarray Analysis	17
<i>Juan F. De Paz, Ana B. Gil, Emilio Corchado</i>	

A Comparative Study of Microarray Data Classification Methods Based on Ensemble Biological Relevant Gene Sets	25
<i>Miguel Reboiro-Jato, Daniel Glez-Peña, Juan Francisco Gálvez, Rosalía Laza Fidalgo, Fernando Díaz, Florentino Fdez-Riverola</i>	

Data Mining and Data Integration

Predicting the Start of Protein α-Helices Using Machine Learning Algorithms	33
<i>Rui Camacho, Rita Ferreira, Natacha Rosa, Vânia Guimarães, Nuno A. Fonseca, Vítor Santos Costa, Miguel de Sousa, Alexandre Magalhães</i>	

A Data Mining Approach for the Detection of High-Risk Breast Cancer Groups	43
<i>Orlando Anunciação, Bruno C. Gomes, Susana Vinga, Jorge Gaspar, Arlindo L. Oliveira, José Rueff</i>	
GRASP for Instance Selection in Medical Data Sets	53
<i>Alfonso Fernández, Abraham Duarte, Rosa Hernández, Ángel Sánchez</i>	
Expanding Gene-Based PubMed Queries	61
<i>Sérgio Matos, Joel P. Arrais, José Luis Oliveira</i>	
Improving Cross Mapping in Biomedical Databases	69
<i>Joel Arrais, João E. Pereira, Pedro Lopes, Sérgio Matos, José Luis Oliveira</i>	
An Efficient Multi-class Support Vector Machine Classifier for Protein Fold Recognition	77
<i>Wiesław Chmielnicki, Katarzyna Stępor, Irena Roterman-Konieczna</i>	
Feature Selection Using Multi-Objective Evolutionary Algorithms: Application to Cardiac SPECT Diagnosis	85
<i>António Gaspar-Cunha</i>	
Phylogenetics and Sequence Analysis	
Two Results on Distances for Phylogenetic Networks	93
<i>Gabriel Cardona, Mercè Llabrés, Francesc Rosselló</i>	
Cramér Coefficient in Genome Evolution	101
<i>Vera Afreixo, Adelaide Freitas</i>	
An Application for Studying Tandem Repeats in Orthologous Genes	109
<i>José Paulo Lousado, José Luis Oliveira, Gabriela Moura, Manuel A.S. Santos</i>	
Accurate Selection of Models of Protein Evolution	117
<i>Mateus Patricio, Federico Abascal, Rafael Zardoya, David Posada</i>	
Scalable Phylogenetics through Input Preprocessing	123
<i>Roberto Blanco, Elvira Mayordomo, Esther Montes, Rafael Mayo, Angelines Alberto</i>	
The Median of the Distance between Two Leaves in a Phylogenetic Tree	131
<i>Arnau Mir, Francesc Rosselló</i>	

***In Silico* AFLP: An Application to Assess What Is Needed to Resolve a Phylogeny** 137
María Jesús García-Pereira, Armando Caballero, Humberto Quesada

Employing Compact Intra-Genomic Language Models to Predict Genomic Sequences and Characterize Their Entropy 143
Sérgio Deusdado, Paulo Carvalho

Biomedical Applications

Structure Based Design of Potential Inhibitors of Steroid Sulfatase 151
Elisangela V. Costa, M. Emília Sousa, J. Rocha, Carlos A. Montanari, M. Madalena Pinto

Agent-Based Model of the Endocrine Pancreas and Interaction with Innate Immune System 157
Ignacio V. Martínez Espinosa, Enrique J. Gómez Aguilera, María E. Hernando Pérez, Ricardo Villares, José Mario Mellado García

State-of-the-Art Genetic Programming for Predicting Human Oral Bioavailability of Drugs 165
Sara Silva, Leonardo Vanneschi

Pharmacophore-Based Screening as a Clue for the Discovery of New P-Glycoprotein Inhibitors 175
Andreia Palmeira, Freddy Rodrigues, Emília Sousa, Madalena Pinto, M. Helena Vasconcelos, Miguel X. Fernandes

Bioinformatics Applications

***e*-BiMotif: Combining Sequence Alignment and Biclustering to Unravel Structured Motifs** 181
Joana P. Gonçalves, Sara C. Madeira

Applying a Metabolic Footprinting Approach to Characterize the Impact of the Recombinant Protein Production in *Escherichia Coli* 193
Sónia Carneiro, Silas G. Villas-Bôas, Isabel Rocha, Eugénio C. Ferreira

Rbbt: A Framework for Fast Bioinformatics Development with Ruby 201
Miguel Vázquez, Rubén Nogales, Pedro Carmona, Alberto Pascual, Juan Pavón

Analysis of the Effect of Reversibility Constraints on the Predictions of Genome-Scale Metabolic Models	209
<i>José P. Faria, Miguel Rocha, Rick L. Stevens, Christopher S. Henry</i>	
Enhancing Elementary Flux Modes Analysis Using Filtering Techniques in an Integrated Environment	217
<i>Paulo Maia, Marcellinus Pont, Jean-François Tomb, Isabel Rocha, Miguel Rocha</i>	
Genome Visualization in Space	225
<i>Leandro S. Marcolino, Bráulio R.G.M. Couto, Marcos A. dos Santos</i>	
A Hybrid Scheme to Solve the Protein Structure Prediction Problem	233
<i>José C. Calvo, Julio Ortega, Mancia Anguita</i>	
Author Index	241