

Studies in Computational Intelligence, Volume 367

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. 346. Weisi Lin, Dacheng Tao, Janusz Kacprzyk, Zhu Li, Ebroul Izquierdo, and Haohong Wang (Eds.)
Multimedia Analysis, Processing and Communications, 2011
ISBN 978-3-642-19550-1

Vol. 347. Sven Helmer, Alexandra Poulouvassilis, and Fatos Xhafa
Reasoning in Event-Based Distributed Systems, 2011
ISBN 978-3-642-19723-9

Vol. 348. Beniamino Murgante, Giuseppe Borruso, and Alessandra Lapucci (Eds.)
Geocomputation, Sustainability and Environmental Planning, 2011
ISBN 978-3-642-19732-1

Vol. 349. Vitor R. Carvalho
Modeling Intention in Email, 2011
ISBN 978-3-642-19955-4

Vol. 350. Thanasis Daradoumis, Santi Caballé, Angel A. Juan, and Fatos Xhafa (Eds.)
Technology-Enhanced Systems and Tools for Collaborative Learning Scaffolding, 2011
ISBN 978-3-642-19813-7

Vol. 351. Ngoc Thanh Nguyen, Bogdan Trawiński, and Jason J. Jung (Eds.)
New Challenges for Intelligent Information and Database Systems, 2011
ISBN 978-3-642-19952-3

Vol. 352. Nik Bessis and Fatos Xhafa (Eds.)
Next Generation Data Technologies for Collective Computational Intelligence, 2011
ISBN 978-3-642-20343-5

Vol. 353. Igor Aizenberg
Complex-Valued Neural Networks with Multi-Valued Neurons, 2011
ISBN 978-3-642-20352-7

Vol. 354. Ljupco Kocarev and Shiguo Lian (Eds.)
Chaos-Based Cryptography, 2011
ISBN 978-3-642-20541-5

Vol. 355. Yan Meng and Yaochu Jin (Eds.)
Bio-Inspired Self-Organizing Robotic Systems, 2011
ISBN 978-3-642-20759-4

Vol. 356. Sławomir Koziel and Xin-She Yang (Eds.)
Computational Optimization, Methods and Algorithms, 2011
ISBN 978-3-642-20858-4

Vol. 357. Nadia Nedjah, Leandro Santos Coelho, Viviana Cocco Mariani, and Luiza de Macedo Mourelle (Eds.)
Innovative Computing Methods and their Applications to Engineering Problems, 2011
ISBN 978-3-642-20957-4

Vol. 358. Norbert Jankowski, Włodzisław Duch, and Krzysztof Grańbczewski (Eds.)
Meta-Learning in Computational Intelligence, 2011
ISBN 978-3-642-20979-6

Vol. 359. Xin-She Yang, and Sławomir Koziel (Eds.)
Computational Optimization and Applications in Engineering and Industry, 2011
ISBN 978-3-642-20985-7

Vol. 360. Mikhail Moshkov and Beata Zielosko
Combinatorial Machine Learning, 2011
ISBN 978-3-642-20994-9

Vol. 361. Vincenzo Pallotta, Alessandro Soro, and Eloisa Vargiu (Eds.)
Advances in Distributed Agent-Based Retrieval Tools, 2011
ISBN 978-3-642-21383-0

Vol. 362. Pascal Bouvry, Horacio González-Vélez, and Joanna Kolodziej (Eds.)
Intelligent Decision Systems in Large-Scale Distributed Environments, 2011
ISBN 978-3-642-21270-3

Vol. 363. Kishan G. Mehrotra, Chilukuri Mohan, Jae C. Oh, Pramod K. Varshney, and Moonis Ali (Eds.)
Developing Concepts in Applied Intelligence, 2011
ISBN 978-3-642-21331-1

Vol. 364. Roger Lee (Ed.)
Computer and Information Science, 2011
ISBN 978-3-642-21377-9

Vol. 365. Roger Lee (Ed.)
Computers, Networks, Systems, and Industrial Engineering 2011, 2011
ISBN 978-3-642-21374-8

Vol. 366. Mario Köppen, Gerald Schaefer, and Ajith Abraham (Eds.)
Intelligent Computational Optimization in Engineering, 2011
ISBN 978-3-642-21704-3

Vol. 367. Gabriel Luque and Enrique Alba
Parallel Genetic Algorithms, 2011
ISBN 978-3-642-22083-8

Gabriel Luque and Enrique Alba

Parallel Genetic Algorithms

Theory and Real World Applications

Authors

Dr. Gabriel Luque
E.T.S.I. Informática
University of Málaga
Campus de Teatinos
29071 Málaga
Spain
Email: gabriel@lcc.uma.es

Prof. Enrique Alba
E.T.S.I. Informática
University of Málaga
Campus de Teatinos
29071 Málaga
Spain
Email: eat@lcc.uma.es

ISBN 978-3-642-22083-8

e-ISBN 978-3-642-22084-5

DOI 10.1007/978-3-642-22084-5

Studies in Computational Intelligence

ISSN 1860-949X

Library of Congress Control Number: 2011930858

© 2011 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 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.

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

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

To my family
Gabriel Luque

To my family, for their
continuous support
Enrique Alba

Preface

This book is the result of several years of research trying to better characterize parallel genetic algorithms (pGAs) as a powerful tool for optimization, search, and learning.

We here offer a presentation structured in three parts. The first one is targeted to the algorithms themselves, discussing their components, the physical parallelism, and best practices in using and evaluating them.

A second part deals with theoretical results relevant to the research with pGAs. Here we stress several issues related to actual and common pGAs.

A final third part offers a very wide study of pGAs as problem solvers, addressing domains such as natural language processing, circuits design, scheduling, and genomics. With such a diverse analysis, we intend to show the big success of these techniques in Science and Industry.

We hope this book will be helpful both for researchers and practitioners. The first part shows pGAs to either beginners or researchers looking for a unified view of the field. The second part partially solves (and also opens) new investigation lines in theory of pGAs. The third part can be accessed independently for readers interested in those applications. A small note on MALLBA, one of our software libraries for parallel GAs is also included to ease laboratory practices and actual applications.

We hope the reader will enjoy the contents as much we did in writing this book.

Málaga,
May 2010

Gabriel Luque
Enrique Alba

Contents

Part I: Introduction

1	Introduction	3
1.1	Optimization	4
1.2	Metaheuristics	5
1.3	Evolutionary Algorithms	7
1.4	Decentralized Genetic Algorithms	10
1.5	Conclusions	12
2	Parallel Models for Genetic Algorithms	15
2.1	Panmictic Genetic Algorithms	17
2.2	Structured Genetic Algorithms	18
2.3	Parallel Genetic Algorithms	19
2.3.1	Parallel Models	20
2.3.2	A Brief Survey on Parallel GAs	23
2.3.3	New Trends in pGAs	25
2.4	First Experimental Results	26
2.4.1	MAXSAT Problem	26
2.4.2	Analysis of Results	27
2.5	Summary	29
3	Best Practices in Reporting Results with Parallel Genetic Algorithms	31
3.1	Parallel Performance Measures	32
3.1.1	Speedup	32
3.1.2	Other Parallel Measures	36
3.2	How to Report Results in pGAs	37
3.2.1	Experimentation	37
3.2.2	Measuring Performance	39
3.2.3	Quality of the Solutions	39
3.2.4	Computational Effort	40

3.2.5	Statistical Analysis	41
3.2.6	Reporting Results	42
3.3	Inadequate Utilization of Parallel Metrics	43
3.4	Illustrating the Influence of Measures	45
3.4.1	Example 1: On the Absence of Information	46
3.4.2	Example 2: Relaxing the Optimum Helps	47
3.4.3	Example 3: Clear Conclusions do Exist	48
3.4.4	Example 4: Meaningfulness Does Not Mean Clear Superiority	48
3.4.5	Example 5: Speedup: Avoid Comparing Apples against Oranges	49
3.4.6	Example 6: A Predefined Effort Could Hinder Clear Conclusions	50
3.5	Conclusions	51

Part II: Characterization of Parallel Genetic Algorithms

4	Theoretical Models of Selection Pressure for Distributed GAs	55
4.1	Existing Theoretical Models	57
4.1.1	The Logistic Model	57
4.1.2	The Hypergraph Model	57
4.1.3	Other Models	58
4.2	Analyzed Models	59
4.3	Effects of the Migration Policy on the Actual Growth Curves	60
4.3.1	Parameters	61
4.3.2	Migration Topology	61
4.3.3	Migration Frequency	63
4.3.4	Migration Rate	64
4.3.5	Analysis of the Results	65
4.4	Takeover Time Analysis	68
4.5	Conclusions	71

Part III: Applications of Parallel Genetic Algorithms

5	Natural Language Tagging with Parallel Genetic Algorithms	75
5.1	Statistical Tagging	77
5.2	Automatic Tagging with Metaheuristics	79
5.2.1	Genetic Algorithm	79
5.2.2	CHC Algorithm	80
5.2.3	Simulated Annealing	80
5.2.4	Parallel Versions	80

5.3	Algorithm Decisions: Representation, Evaluation, and Operators	81
5.3.1	Individuals	81
5.3.2	Fitness Evaluation	82
5.3.3	Genetic Operators	82
5.4	Experimental Design and Analysis	83
5.5	Conclusions	89
6	Design of Combinational Logic Circuits	91
6.1	Problem Definition	92
6.2	Encoding Solutions into Strings	94
6.3	Related Works	97
6.4	Sequential, Parallel, and Hybrid Approaches	97
6.5	Computational Experiments and Analysis of Their Results	102
6.5.1	Case Study 1: Sasao	104
6.5.2	Case Study 2: Catherine	106
6.5.3	Case Study 3: Katz 1	108
6.5.4	Case Study 4: 2-Bit Multiplier	109
6.5.5	Case Study 5: Katz 2	110
6.6	Overall Discussion	113
6.7	Conclusions and Future Work	114
7	Parallel Genetic Algorithm for the Workforce Planning Problem	115
7.1	The Workforce Planning Problem	116
7.2	Design of a Genetic Algorithm	118
7.2.1	Solution Encoding	118
7.2.2	Evaluation the Quality of a Solution	119
7.2.3	Repairing/Improving Operator	120
7.2.4	Recombination Operator	120
7.2.5	Mutation Operator	122
7.2.6	The Proposed Parallel GA	122
7.3	Scatter Search	123
7.3.1	Seeding the Initial Population	124
7.3.2	Improvement Method	124
7.3.3	Parallel SS	125
7.4	Computational Experiments and Analysis of Results	125
7.4.1	Problem Instances	126
7.4.2	Results: Workforce Planning Performance	126
7.4.3	Results: Computational Times	129
7.4.4	A Parallel Hybrid GA	132
7.5	Conclusions	134

8	Parallel GAs in Bioinformatics: Assembling DNA	
	Fragments	135
8.1	The Work of a DNA Fragment Assembler	136
8.1.1	DNA Sequencing Process	136
8.2	Related Literature	139
8.3	The pGA DNA Assembler	140
8.3.1	Solution Encoding	140
8.3.2	Solution Evaluation	140
8.3.3	Genetic Operators	141
8.3.4	The Parallel Approach	142
8.4	Experimental Validation	143
8.4.1	Target Problem Instances	143
8.4.2	Parameterization	144
8.4.3	Analysis of Results	144
8.5	Conclusions	147
A	The MALLBA Library	149
A.1	Skeleton Interfaces	151
A.2	Communication Interface	152
A.3	Hybridization Interface	154
A.4	Additional Information about MALLBA	156
B	Acronyms	157
	References	159