Natural Computing Series

Series Editors: G. Rozenberg
Th. Bäck A.E. Eiben J.N. Kok H.P. Spaink

Leiden Center for Natural Computing

For further volumes:
http://www.springer.com/series/4190

Yossi Borenstein • Alberto Moraglio
Editors

Theory and Principled Methods for the Design of Metaheuristics

Springer

Editors
Yossi Borenstein
VisualDNA
London
United Kingdom

Alberto Moraglio
University of Birmingham
School of Computer Science
Birmingham
United Kingdom

ISSN 1619-7127 Natural Computing Series
ISBN 978-3-642-33205-0 ISBN 978-3-642-33206-7 (eBook)
DOI 10.1007/978-3-642-33206-7
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013956610

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

For Aloe, Annette, and Shlomo—Y.B.

For Fumiyo, Giuseppina, and Giuseppe—A.M.

Foreword

I was very happy and excited when I was invited to write a foreword for the book "Theory and Principled Methods for the Design of Metaheuristics" by Alberto Moraglio and Yossi Borenstein. There are several reasons for this.

Firstly, the focus of the book—bridging the gap between theory and practice—is an exceptionally important and timely topic in the field of metaheuristics and is one that is very dear to me. Despite numerous attempts to change this, by and large theoreticians and practitioners live on different planets. Often theoretical studies on metaheuristics are hardly applicable to real-world problems: so why should practitioners take notice? Conversely, understanding theoretical results requires some discipline and effort, and I must say that practitioners, often unjustifiably, do not think that what they can learn from theory is worth the investment of their time, when in fact it is. My feeling is that this book will be an important reference in this area from which both theoreticians and practitioners can learn much.

Secondly, I was really impressed with the list of contributors to this edited book. This includes many of the most influential and respected researchers on metaheuristics. This book is a fantastic fusion of their collective knowledge. In terms of both contributions to the science and engineering of metaheuristics and inspiration for future researchers and practitioners, this book appears to have been a resounding success.

Finally, not too long ago, both Alberto and Yossi were doctoral students under my supervision. They were both exceptionally good students—independent thinkers with sharp minds and a strong motivation and will. It is both very satisfying and a source of inspiration for me to see them actively working with a number of world-class researchers on casting light on some of the most difficult yet crucially important issues in metaheuristics and much beyond, at a time when my own will is faltering.

I am sure I will learn a lot from studying this volume in detail, and I am sure many others, practitioners and theoreticians alike, can gain much from doing the same—Yossi and Alberto, thank you for having put together such a high-profile and inspiring volume.

Colchester, Essex, UK — Riccardo Poli

Preface

Metaheuristics and evolutionary algorithms in particular are adaptable optimization frameworks that are routinely and successfully applied to hard real-world problems. Intuitively, they seem to capture some fundamental biological property which makes them inherently good general problem solvers. At the same time, the informal way in which metaheuristics are defined and tailored to each problem can lead to various misconceptions. More than anything else, their successful application is often the outcome of a long trial-and-error process to identify a good problem-specific design followed by extensive parameter tuning. Ideally, theoretical studies should rectify this situation by explaining how, when, and why metaheuristics work and providing guidelines for their successful design and optimal parameter choice. However, the challenge is huge: mathematical analysis requires time and effort even for very simple scenarios, whereas in practice problems as well as algorithms are quite complex and subject to rapid change. The different motivations—practical applications vs. mathematical analysis—lead to distinct subcommunities with different research cultures which rarely communicate with one another.

The dialog between theory and practice is very important to us. We organized a workshop at Parallel Problem Solving from Nature (PPSN) on this subject and later edited a special issue of the Evolutionary Computation Journal (ECJ). Our study is theoretical, but we always made an effort to identify and promote potential practical applications. In this book, rather than focusing on our own work, we collected several theoretical and principled methods that when strung together we believe indicate a viable route towards bridging theory and practice. The book outlines the contribution of current theoretical work to practical problems, points to various theoretical approaches that have the potential to have a real impact on practice in the future, and, at the same time, provides principled methods that can be applied now to make empirical work more rigorous. It is divided into four themes: the first three are related to theory and the last one to practice.

The Story Told in the Book

The wildest dream of those working with metaheuristics is to have a single universal search algorithm that could be applied unchanged to any problem and that would always deliver the optimal solution, efficiently. Claims along this line have been made in the past about genetic algorithms. It is clear, nowadays, that such an algorithm does not exist. One of the roles of theory is to defy claims of the sort. In the section "Theory for Drawing the Line," we give two examples of theoretical results that show what is *not* possible.

Results such as the no-free-lunch theorems show that it is necessary to make a compromise between the class of problems that a search algorithm is applied to and its overall expected performance. The most common counterargument to no-free-lunch theorems is that problems of practical interest are a very small subclass of all possible problems, and commonly used metaheuristics do well exactly on this class. This argument, however, begs the question of what really accounts for an interesting problem. In the section "Relevant Scope of Problems," we give three examples for possible ways of defining formally general classes of real-world problems.

The requirement to match problem class and search algorithm can be also looked at the other way around. Given a not too large rigorously defined class of problems, in principle it could be possible to design a search algorithm that is provably good for this class. In the section "Top-Down Principled Design of Search Algorithms," we give three examples of works that pursue this line of investigation.

The outline of a theory given above (i.e., formally defining an interesting general class of problems and then, accordingly, developing an optimal search algorithm for this class) has the potential to be the ultimate tool for practitioners. In principle, once the practitioner identifies a problem as a specific case of a more general class, he/she will have a choice of different optimal search algorithms designed for that class with guaranteed expected optimization time. Unfortunately, it is very challenging to find a balance between a class of problems which is broad enough to be practically interesting and yet is focused enough to admit an efficient search algorithm. Therefore, it is difficult to estimate when such a vision will become a practicable reality.

For the time being, it is therefore necessary to embrace an experimental approach to the application of metaheuristics to specific problems. Nonetheless, existing theory can be a guide for good practice. The section "Principled Practice" is about reasoned and systematic approaches to setting up experiments, metaheuristic adaptation to the problem at hand, and parameter settings. We give three examples of such works.

Overview of the Chapters

Theory for Drawing the Line

Knowing what is not possible avoids tempting but hopeless lines of research. The first two contributions present theoretical results that were developed as a response to empirical attempts to chase chimeras.

In the first chapter, “No Free Lunch Theorems: Limitations and Perspectives of Metaheuristics,” Christian Igel reviews the no-free-lunch theorems for search and optimization, and their implications for the design of metaheuristics are discussed. The no-free-lunch theorems show that it is not possible to develop a black-box search algorithm that is universally better than any other on every problem. Search algorithms must be tailored to the problem class at hand using prior knowledge to deliver good performance.

Fabien Teytaud and Olivier Teytaud consider in the second chapter, “Convergence Rates of Evolutionary Algorithms and Parallel Evolutionary Algorithms,” a large family of search algorithms that uses comparisons rather than absolute fitness values in the selection process. The focus on comparisons—even without considering specific classes of problems—is sufficient to demonstrate advantages in terms of robustness and, at the same time, drawbacks in terms of diminished performance. Practical implications of these results for evolutionary algorithms on parallel machines are discussed.

Relevant Scope of Problems

In order to design a “better than random search” algorithm, it is necessary to restrict the scope of the problems one considers. If the scope is too large, the gain in performance may be not practically relevant. If the scope is too narrow, the search algorithm may not be of any general interest (other than to the very specific problem at hand). If the scope excludes real-world problems, it will not be interesting, even if it encompasses a fairly large number of problems and it works substantially better than other algorithms on this class of problems. The scope needs to be defined rigorously; this will make it possible to: avoid improper claims, be a starting point for devising search algorithms matching it, and serve as a starting point to prove general results (on the performance of search algorithms) on this class of problems. The following three contributions describe research attempting to identify interesting classes of problems.

In Chap. 3, “Rugged and Elementary Landscapes,” Konstantin Klemm and Peter F. Stadler provide an introduction to the structural features of discrete fitness landscapes from both the geometric and the algebraic perspectives. In particular, the chapter focuses on elementary landscapes, which are a class of fitness landscapes that encompass several important real-world problems.

In Chap. 4, “Single-Funnel and Multi-funnel Landscapes and Subthreshold-Seeking Behavior,” Darrell Whitley and Jonathan Rowe introduce the classes of single-funnel and multi-funnel landscapes. These classes of problems are quite large; however, they capture the characteristics of many typical real-world problems. They show that a simple subthreshold-seeker algorithm performs provably better than random search on these classes.

Chapter 3 introduces an important class of problems but without devising a better than random search algorithm for that class. Chapter 4 provides such an algorithm for a very large class of problems; however, the size of this class limits the potential performance of the algorithm. In Chap. 5 “Black-Box Complexity for Bounding the Performance of Randomized Search Heuristics,” Thomas Jansen considers more specific classes of problems and provides optimal randomized search heuristics for those problems. This chapter highlights the importance of focusing on specific classes of problems. It also exemplifies, using the notion of black-box complexity, how one can theoretically prove optimality for black-box algorithms (and hence make any attempts to design better algorithms redundant).

Top-Down Principled Design of Search Algorithms

The features of the class of problems considered can be used to derive in a principled way search algorithms that, exploiting these properties, reach the best possible (average) result on the considered class. For example, for a strictly unimodal fitness landscape, we might want to use a steepest-descent local search algorithm. This way, we take advantage of the special feature of this problem—that any local optimum reached from any starting point is the global optimum. The first two contributions illustrate how to derive a search algorithm that is optimally matched in a certain sense with a probabilistic class of functions. Rather than considering an explicit class of problems, the third contribution shows how the well-known Covariance Matrix Adaptation Evolution Strategy (CMA-ES) was derived by exploiting a desirable property of such class.

In Chap. 6, “Designing an Optimal Search Algorithm with Respect to Prior Information,” Olivier Teytaud and Emmanuel Vazquez consider three approaches to derive an optimal search algorithm for a class of functions: experimentation (i.e., parameter tuning), a mathematical approach based on reinforcement learning, and a simplified version of the latter with more reasonable computational cost based on Gaussian processes.

In Chap. 7, “The Bayesian Search Game,” Marc Toussaint draws links between no-free-lunch theorems that, interpreted inversely, lay the foundation of how to design search heuristics that exploit prior knowledge about the function, partially observable Markov decision processes and their approach to the problem of sequentially and optimally choosing search points, and the use of Gaussian processes as a representation of belief, i.e., knowledge about the problem.

In Chap. 8, “Principled Design of Continuous Stochastic Search: From Theory to Practice,” Nikolaus Hansen and Anne Auger derive the well-known *covariance* *matrix adaptation evolution strategy*, which has been shown to work very well in continuous optimization occurring in practice. They show how this algorithm was developed based only on a few fundamental principles—namely, maximal entropy, unbiasedness, maintaining invariance, and, under these constraints, exploiting all available information and solving simple functions reasonably fast.

Principled Practice

The literature is rich with an ever-increasing number of new metaheuristics that have demonstrated, one way or another, their potential usefulness. Albeit, metaheuristics are far from being plug-and-play friendly: Given a problem, one has initially to choose which metaheuristic to use. Then, as metaheuristics are not ready-made search algorithms, it is necessary to spend a considerable amount of time on adapting the operators for the particular problem domain and tuning the parameters. The final three contributions suggest how to address these more practical issues in a reasoned and systematic way.

In Chap. 9, "Parsimony Pressure Made Easy: Solving the Problem of Bloat in GP," Riccardo Poli and Nicholas Freitag McPhee use Price's theorem to characterize mathematically the size evolution of programs and to derive theoretical results that show how to practically and optimally use the parsimony pressure method to achieve complete control over the growth of the programs in a population.

In Chap. 10, "Experimental Analysis of Optimization Algorithms: Tuning and Beyond," Thomas Bartz-Beielstein and Mike Preuss present a tutorial on methodological approaches for experimental research in evolutionary computation and metaheuristic optimization.

In the last contributed chapter entitled "Formal Search Algorithms + Problem Characterizations = Executable Search Strategies," Patrick D. Surry and Nicholas J. Radcliffe present a principled way to derive search operators for nonstandard solution representations which also takes into account the structure of the problem at hand.

London, UK
July 2012

Yossi Borenstein
Alberto Moraglio

Acknowledgements

We would like to express our gratitude to the following referees who gave their time and energy to review the chapters of the book ensuring their quality:

Dimo Brockhoff
Ying-Ping Chen
Marco Chiarandini
Cecilia Di Chio
Tobias Friedrich
Mario Giacobini
Mario Graff
Li Hui
Christian Igel
Thomas Jansen
Colin Johnson
Bryant A. Julstrom
Yong-Hyuk Kim
Dario Landa-Silva
Per Kristian Lehre
Guanzhou Lu
Yusuke Nojima
Pietro Oliveto
Mike Preuss
Ramon Sagarna
Jonathan Shapiro
Terence Soule
Peter F. Stadler
Olivier Teytaud
Marc Toussaint
Carsten Witt
Yourim Yoon

Contents

List of Contributors

Anne Auger INRIA Saclay – Île-de-France, Orsay, France

Thomas Bartz-Beielstein Faculty of Computer Science and Engineering Science, Institute of Computer Science, Cologne University of Applied Sciences, Cologne, Germany

Nikolaus Hansen INRIA Saclay – Île-de-France, Orsay, France

Christian Igel Department of Computer Science, University of Copenhagen, Copenhagen, Denmark

Thomas Jansen Department of Computer Science, Aberystwyth University, Aberystwyth, UK

Konstantin Klemm Bioinformatics Group, Department of Computer Science, Interdisciplinary Center for Bioinformatics, University of Leipzig, Leipzig, Germany

Nicholas Freitag McPhee Division of Science and Mathematics, University of Minnesota, Morris, MN, USA

Riccardo Poli School of Computer Science and Electronic Engineering, University of Essex, Colchester, Essex, UK

Mike Preuss Algorithm Engineering, Department of Computer Science, TU Dortmund, Dortmund, Germany

Nicholas J. Radcliffe Stochastic Solutions Limited, Edinburgh, UK
Department of Mathematics, University of Edinburgh, Edinburgh, UK

Jonathan Rowe Department of Computer Science, University of Birmingham, Birmingham, UK

Peter F. Stadler Bioinformatics Group, Department of Computer Science, Interdisciplinary Center for Bioinformatics, University of Leipzig, Leipzig, Germany
Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

Fraunhofer Institut für Zelltherapie und Immunologie, Leipzig, Germany
Department of Theoretical Chemistry, University of Vienna, Wien, Austria
Santa Fe Institute, Santa Fe, NM, USA

Patrick D. Surry Portrait Software, Boston, MA, USA

Fabien Teytaud TAO, Inria Saclay IDF, LRI, University Paris-Sud, Paris, France

Olivier Teytaud TAO, Inria Saclay IDF, LRI, University Paris-Sud, Paris, France

Marc Toussaint Machine Learning & Robotics Lab, Free University of Berlin, Berlin, Germany

Emmanuel Vazquez Department of Computer Science and Information Engineering, National University of Tainan, Tainan, Taiwan
SUPELEC, Gif-sur-Yvette, France

Darrell Whitley Department of Computer Science, Colorado State University, Fort Collins, CO, USA