Theory of Randomized Optimization Heuristics (Report of Dagstuhl Seminar 19431)

Carola Doerr, Carlos Fonseca, Tobias Friedrich, Xin Yao

Report from Dagstuhl Seminar 19431

Theory of Randomized Optimization Heuristics

Edited by

Carola Doerr[1], Carlos M. Fonseca[2], Tobias Friedrich[3], and Xin Yao[4]

1 **Sorbonne University – Paris, FR,** carola.doerr@mpi-inf.mpg.de
2 **University of Coimbra, PT,** cmfonsec@dei.uc.pt
3 **Hasso Plattner Institute – Potsdam, DE,** friedrich@hpi.de
4 **Southern Univ. of Science and Technology – Shenzen, CN,** xiny@sustech.edu.cn

Abstract

This report documents the activities of Dagstuhl Seminar 19431 on "Theory of Randomized Optimization Heuristics". 46 researchers from Europe, Australia, Asia, and North America have come together to discuss ongoing research. This tenth edition of the seminar series had three focus topics: (1) relation between optimal control and heuristic optimization, (2) benchmarking optimization heuristics, and (3) the interfaces between continuous and discrete optimization. Several breakout sessions have provided ample opportunity to brainstorm on recent developments in the research landscape, to discuss and solve open problems, and to kick-start new research initiatives.

Seminar October 20–25, 2019 – http://www.dagstuhl.de/19431

2012 ACM Subject Classification Theory of computation → Bio-inspired optimization, Theory of computation → Evolutionary algorithms, Theory of computation → Optimization with randomized search heuristics

Keywords and phrases algorithms and complexity, evolutionary algorithms, machine learning, optimization, soft computing

Digital Object Identifier 10.4230/DagRep.9.10.61

Edited in cooperation with Patrick Spettel

1 Executive Summary

Carola Doerr
Carlos M. Fonseca
Tobias Friedrich
Xin Yao

Efficient optimization techniques affect our personal, industrial, and academic environments through the supply of well-designed processes that enable a best-possible use of our limited resources. Despite significant research efforts, most real-world problems remain too complex to admit exact analytical or computational solutions. Therefore, heuristic approaches that trade the accuracy of a solution for a simple algorithmic structure, fast running times, or an otherwise efficient use of computational resources are required. Randomized optimization heuristics form a highly successful and thus frequently applied class of such problem solvers. Among the best-known representatives of this class are stochastic local search methods, Monte Carlo techniques, genetic and evolutionary algorithms, and swarm intelligence techniques.

Theory of Randomized Optimization Heuristics, *Dagstuhl Reports*, Vol. 9, Issue 10, pp. 61–94
Editors: Carola Doerr, Carlos M. Fonseca, Tobias Friedrich, and Xin Yao
Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

The theory of randomized optimization heuristics strives to set heuristic approaches on firm ground by providing a sound mathematical foundation for this important class of algorithms. Key challenges in this research area comprise optimization under uncertainty, parameter selection (most randomized optimization heuristics are parametrized), the role and usefulness of so-called *crossover* operations (i.e., the idea of creating high-quality solution candidates by recombining previously evaluated ones) and, more generally, performance guarantees for advanced heuristics such as population-based techniques, estimation-of-distribution algorithms, differential evolution, and others.

Dagstuhl Seminar 19431 on "Theory of Randomized Optimization Heuristics" was a continuation of the seminar series originally on "Theory of Evolutionary Algorithms". Today the field extends far beyond evolutionary algorithms – a development that previous Dagstuhl seminars have significantly influenced.

While the previous seminar 17191 had a very strong focus on methodological questions and techniques needed to analyze stochastic optimization heuristics, the present seminar had among its three main focus topics chosen to foster interaction with two strongly linked research communities that were not previously represented in the seminar series: stochastic control theory and empirical benchmarking of randomized optimization heuristics.

Recent work has shown that there is a very close link between the theory of randomized optimization heuristics and stochastic control theory, both regarding the nature of the "systems" of interest and the analytical techniques that have been developed in the two communities. At the seminar, we have explored these affinities through the two invited presentations of Luc Pronzato and Vivek Borkar, through contributed talks highlighting different aspects studied in both communities (e.g., the presentation on one-shot optimization by Olivier Teytaud), and through focussed breakout sessions, in particular the one fully dedicated to *Connection between the analysis of evolution strategies and estimation of distribution algorithms and the analysis of stochastic approximation and ordinary differential equations*, in which interesting similarities and differences between the two fields were identified.

The second focus topic of Dagstuhl Seminar 19431 was benchmarking of optimization heuristics. Benchmarking plays a central role in empirical performance assessment. However, it can also be an essential tool for theoreticians to develop their mathematically-derived ideas into practical algorithms, thereby encouraging a principled discussion between empirically-driven and theoretically-driven researchers. Benchmarking has been a central topic in several breakout sessions, for example those on *Competitions and Benchmarking*, *Algorithm Selection and Configuration*, but also the breakout session on *Multi-Objective Optimization.* A survey of best practices in empirical benchmarking has been kick-started in the breakout session on *Benchmarking: Best Practices and Open Issues.*

Discussing the mathematical challenges arising in the performance analysis of randomized heuristics has always been a central topic in this Dagstuhl seminar series. Among other achievements, important connections between continuous and discrete optimization have been established, most notably in the form of drift theorems, which are typically applicable regardless of the nature of the search space. Apart from such methodological advances, we have also observed two other trends bridging discrete and continuous optimization: (i) an increased interest in analyzing parameter-dependent performance guarantees, and (ii) the recent advances in the study of estimation of distribution algorithms, which borrow techniques from both discrete and continuous optimization theory. These topics have been discussed in the invited talk of Youhei Akimoto, in several contributed presentations, and in the breakout sessions on *Measuring Optimization Progress in an Invariant Way for Comparison-Based Algorithms* and on *Mixed-Integer Optimization.*

Apart from these focus topics, we have discussed a large number of different aspects related to the theoretical analysis of optimization heuristics, including brainstorming sessions on doing "good" research, organizing a repository to share lecture materials, and discussing the role of uncertainty in heuristic optimization, the connections between experimental design and one-shot optimization, the importance of neutral representations, and differences between stochastic gradient descent methods and evolution strategies, to give but a few examples.

Organization

The seminar hosted the following type of events:

- Five invited talks of 30 minutes each:
  - Youhei Akimoto on *Expected Runtime Bound for the (1+1)-Evolution Strategy*
  - Vivek Borkar on *Overview of Stochastic Approximation and Related Schemes*
  - Pietro S. Oliveto on *What is Hot in Evolutionary Computation (Part 2)*
  - Luc Pronzato on *Dynamical Search*
  - Carsten Witt on *What is Hot in Evolutionary Computation (Part 1)*
- 20 contributed talks of around 15-20 minutes
- Four "flash talks" of about 10 minutes
- Eleven parallel breakout sessions in various different formats, ranging from brainstorming on the purpose and future of theory research through actual problem solving on one-shot optimization to kick-starting a survey on best practices on benchmarking optimization heuristics.

All presentations were plenary, i.e., in a single session, while the breakouts were organized in parallel working groups, to allow for focused and specialized discussions. As in previous years, the breakout sessions were very well perceived, and can be considered a well-established format of this seminar series. As a result of these discussions, we are planning a workshop and a survey on benchmarking best practices. Several open problems have been proposed and discussed at the seminar, and we are confident that the seminar has helped to establish new collaborations.

Our traditional hike on Wednesday was a good opportunity to discuss in a less formal setting and to get to know each other. On Thursday evening, we had the special opportunity to hear Jonathan Rowe present activities of the Alain Turing Institute https://www.turing.ac.uk/, where he serves as Programme Director for Data Science for Science. Last, but not least, the wine-and-cheese party complemented the scientific activities with a relaxed social event.

We would like to thank the Dagstuhl team and all participants for making seminar 19431 a great success and a great pleasure to organize.

Carola Doerr (Sorbonne University – Paris, FR)
Carlos M. Fonseca (University of Coimbra, PT)
Tobias Friedrich (Hasso Plattner Institute – Potsdam, DE)
Xin Yao (Southern University of Science and Technology – Shenzen, CN)

2 Table of Contents

3 Overview of Talks

3.1 Expected Runtime Bounds for $(1+1)$-ES

Youhei Akimoto (University of Tsukuba, JP)

Main reference Daiki Morinaga, Youhei Akimoto: "Generalized drift analysis in continuous domain: linear convergence of (1 + 1)-ES on strongly convex functions with Lipschitz continuous gradients", in Proc. of the 15th ACM/SIGEVO Conference on Foundations of Genetic Algorithms, FOGA 2019, Potsdam, Germany, August 27-29, 2019, pp. 13–24, ACM, 2019.
URL https://doi.org/10.1145/3299904.3340303
Main reference Youhei Akimoto, Anne Auger, Tobias Glasmachers: "Drift theory in continuous search spaces: expected hitting time of the (1 + 1)-ES with 1/5 success rule", in Proc. of the Genetic and Evolutionary Computation Conference, GECCO 2018, Kyoto, Japan, July 15-19, 2018, pp. 801–808, 2018.
URL https://doi.org/10.1145/3205455.3205606

We presented recent results on the expected runtime bound for (1+1)-ES by drift analysis. We focused on what we want to show, what has been done, and what are still open. In the end, we had a discussion on what are the similarities and dissimilarities between drift analysis in continuous domain and discrete domain.

3.2 Precise Analysis for Plateaus

Denis Antipov (ITMO University – St. Petersburg, RU) and Benjamin Doerr (Ecole Polytechnique – Palaiseau, FR)

Main reference Denis Antipov, Benjamin Doerr: "Precise Runtime Analysis for Plateaus", CoRR, Vol. abs/1806.01331, 2018.
URL https://arxiv.org/abs/1806.01331

Local optima and plateaus are the features of the fitness landscape which usually make a fitness function hard to be optimized by random search heuristics. While there are plenty of works considering the problem of escaping local optima, most of which are based on the jump functions, plateaus have not got this much attention in the community. In this talk we consider our results on analysis of the simple mutation-based EAs on a benchmark $\textsc{Plateau}_k$ function, introduce the new methods of the analysis for the plateaus and discuss what are the obstacles for spreading these methods on the more complex algorithms.

3.3 A Unified Invariance Formalism for Discrete and Continuous Optimization

Anne Auger (INRIA Saclay – Palaiseau, FR)

Invariance is a general concept that is fundamental in many domains. In statistics, machine learning, decisions taken as a result of a procedure/algorithm based on data should not be affected by simple transformations on the input data like reordering or translation. Invariance is also essential in optimization where we do not want the performance of an algorithm to be greatly affected if e.g. the function optimized is translated or scaled by a positive factor. In this talk I will give a (unified) definition of invariance in the search space that holds in particular for discrete and continuous domains.

3.4 Evolution Strategies are NOT Gradient Followers

Hans-Georg Beyer (Fachhochschule Vorarlberg – Dornbirn, AT)

This talk addresses the question how Evolution Strategies (ES) explore high-dimensional $\mathbb{R}^N$ search spaces. A sometimes invoked picture tries to explain the working by some kind of gradient following strategy. On the other hand there are optimization algorithms that are labeled as ES, however, are actually gradient approximation strategies. It is shown that one of these novel "ESs" resembles the well-known Evolutionary Gradient Search strategy, published in the late 1990s by R. Salomon. Coming back to the question whether classical ESs are gradient approximating strategies, it is shown that this picture does not hold, neither when considering the search behavior of the population in the search space nor when investigating the mean value dynamics of the search process. It turns out that the ES devotes only small steps toward the optimizer in the search space while performing large step in the perpendicular $(N-1)$-dimensional subspace. The one-dimensional part, responsible for the fitness improvement, may be regarded as the exploitation part of the search process while the step in the perpendicular subspace may be regarded as exploration. The ratio of these two steps is roughly proportional to $1/\sqrt{N}$. This is different from vanilla gradient strategies and might explain in parts the success of ESs in highly multimodal optimization problems.

3.5 Stochastic Approximation: an Overview

Vivek Shripad Borkar (Indian Institute of Technology Bombay, IN)

The talk introduces the Robbins-Monro "stochastic approximation" algorithm and the "o.d.e." (for "ordinary differential equations") approach for its convergence analysis. Other theoretical issues such as avoidance of unstable equilibria, limit theorems, stability of iterates, etc. are briefly discussed. Variants such as constant stepsize, two time scale schemes, Markov noise, differential inclusion limits, and distributed asynchronous schemes are mentioned. As example, stochastic gradient and gradient-like schemes are presented. Finally, consensus algorithms are briefly discussed.

3.6 Variations on the Theme of the $(1 + (\lambda, \lambda))$ GA

Maxim Buzdalov (ITMO University – St. Petersburg, RU)

Joint work of Maxim Buzdalov, Anton Bassin

The $(1 + (\lambda, \lambda))$ genetic algorithm is an interesting theory-driven algorithm with many good properties, e.g. the $O(n)$ runtime on OneMax and the showcase of self-adjusting parameter tuning being a crucial part. However, it is still quite slow to conquer other territories, e.g. it is not a very brilliant player even for linear functions. This talk presents our preliminary work on changing this situation. In particular, we introduce a rather successful extension of the $(1 + (\lambda, \lambda))$ GA on problems defined on permutations, and show a few interesting consequences of that regarding how to understand the driving forces behind this algorithm. Another orthogonal idea is that the selection in this algorithm may be rethought based on statistical ideas. This instantly leads to the $O(n)$ runtime on the BinVal function and might pave the road towards wider applicability of this wonderful algorithm.

3.7 Challenges of Mutation Rate Control in $(1 + \lambda)$ Evolutionary Algorithm

Arina Buzdalova (ITMO University – St. Petersburg, RU)

Joint work of Arina Buzdalova, Kirill Antonov, Maxim Buzdalov, Carola Doerr, Irina Petrova, Anna Rodionova
Main reference Anna Rodionova, Kirill Antonov, Arina Buzdalova, Carola Doerr: "Offspring population size matters when comparing evolutionary algorithms with self-adjusting mutation rates", in Proc. of the Genetic and Evolutionary Computation Conference, GECCO 2019, Prague, Czech Republic, July 13-17, 2019, pp. 855–863, 2019.
URL https://doi.org/10.1145/3321707.3321827

It was empirically observed that efficiency of mutation rate control in $(1 + \lambda)$ EA depends on the specified lower bound. Particularly, with the growth of population size λ a higher mutation rate bound of $1/n$ is more efficient than $1/n^2$. However, it seems sensible that a successful adjustment mechanism should not be harmed by a more generous lower bound. We propose a simple modification that makes the 2-rate $(1 + \lambda)$ EA [1] much less sensitive to the lower bound. The open question is how to capture this improvement by theoretical analysis.

References

1 B. Doerr, C. Gießen, C. Witt, and J. Yang, "The $(1 + \lambda)$ Evolutionary Algorithm with Self-Adjusting Mutation Rate", Algorithmica 81, 2 (2019), 593–631.

3.8 Dynastic Potential Crossover Operator

Francisco Chicano (University of Málaga, ES)

An optimal recombination operator provides an optimal solution fulfilling the gene transmission property: the value of any variable in the offspring must be inherited from one of the parents. In the case of binary variables, the offspring of an optimal recombination operator is optimal in the smallest hyperplane containing the two parent solutions. In general, exploring this hyperplane is computationally costly, but if the objective function has a low number of nonlinear interactions among the variables, the exploration can be done in $O(4^{\beta}(n+m)+n^2)$ time, for problems with n decision variables, m subfunctions composing the objective function and where β is a constant. In this talk, we present a quasi-optimal recombination operator, called Dynastic Potential Crossover (DPX), that runs in $O(4^{\beta}(n+m)+n^2)$ time in any case and is able to act as an optimal recombination operator for low-epistasis combinatorial problems. We show some experimental results where the operator is integrated in DRILS (an ILS with recombination) and standard EA solving NKQ Landscapes and MAX-SAT.

3.9 Adaptation of a Sampling Distribution for Metropolis-Hastings

Alexandre Chotard (Calais University, FR)

A Metropolis-Hastings algorithm aims to emulate sampling from a target probability distribution π by using a proposal distribution q. As in optimization, one may adapt the proposal distribution using the points sampled so far, but one also has to care not to bias the resulting stationary distribution.

3.10 Genetic Drift in EDAs

Benjamin Doerr (Ecole Polytechnique – Palaiseau, FR)

Joint work of Benjamin Doerr, Weijie Zheng
Main reference Benjamin Doerr, Weijie Zheng: "Sharp Bounds for Genetic Drift in EDAs", CoRR, Vol. abs/1910.14389, 2019.
URL https://arxiv.org/abs/1910.14389

It has been observed in various mathematical runtime analyses of estimation-of-distribution algorithms that also in the complete absence of a fitness signal, the sampling distributions of the solution values develop a strong preferences for a single value. In this work, we quantify precisely this so-called genetic drift for the univariate EDAs cGA and PBIL (which includes the UMDA and the MMAS ant colony optimizer). Our results suggest how to choose the parameters of these algorithms such as to avoid genetic drift, which is useful both in applications and in research.

3.11 On Potential for Transfer of Results from Theory of Evolutionary Algorithms to Biology

Anton V. Eremeev (Institute of Scientific Information for Social Sciences RAS – Moscow, RU)

The well-known biotechnological procedure SELEX (Systematic Evolution of Ligands by EXponential enrichment) is considered as an experimental implementation of an evolutionary algorithm (EA). As a proof of concept, theoretical bounds on the expected EA runtime and on fraction of sufficiently fit individuals in population are applied in order to forecast the efficiency of SELEX in searching for a promoter sequence, including an enhancer. A comparison of theoretical bounds to the results of computational simulation indicates some cases where the theoretical runtime bounds and bounds on the frequency of highly fit individuals give favorable prediction, while simulation requires prohibitive computational resource. It is shown that further research is needed to extend applicability of the theoretical bounds for the expected runtime and to improve their tightness.

3.12 A (General) Definition of Invariance

Nikolaus Hansen (INRIA Saclay – Palaiseau, FR)

Invariance is arguably one of the single most important conceptions in science. Here, we attempt to give a concise definition of invariance in the context of randomized search algorithms.

▶ **Definition 1** (General Invariance). *Let $\mathcal{F}$ be the set of all functions on a given search space and $\mathcal{H}$ a mapping of $\mathcal{F}$ into its power set,*

$$\mathcal{H} : f \mapsto \mathcal{H}(f) \subset \mathcal{F} \text{ and } f \in \mathcal{H}(f) \text{ w.l.o.g.}\ .$$

We say that a search algorithm is invariant under $\mathcal{H}$ if for every pair of functions $f, h \in \mathcal{H}(f)$ there exists

- *a bijective search space transformation $\varphi_{f\to h}$ and*
- *for all (initial) algorithm states θ on f*
  - *a reachable (initial) state θ' on h and*
  - *a coupling for the random input*

such that for all time steps t the evaluated solutions on h and f, that is the search traces, are equivalent in that

$$x_t^{\theta',h} = \varphi_{f\to h}\left(x_t^{\theta,f}\right)\ .$$

3.13 Gradient Descent and Evolution Strategies are Almost the Same (but don't behave almost the same)

Nikolaus Hansen (INRIA Saclay – Palaiseau, FR)

Modifying the update equations of the iterate in gradient descent to become the update equations of the mean in evolution strategies requires only three surprisingly small changes when given a suitable representation. This insight allows to concisely scrutinize the possible reasons why (and when) gradient descent and evolution strategies behave very differently.

3.14 The UMDA on LeadingOnes Revisited

Martin S. Krejca (Hasso Plattner Institute – Potsdam, DE)

This talk will showcase some of the joint and ongoing work with Benjamin Doerr. When considering the univariate marginal distribution algorithm (UMDA) in a parameter regime with low genetic drift, it can be easily analyzed on the LeadingOnes function. This simplified analysis also improves the currently best known run time bound of the UMDA for that parameter regime.

3.15 Runtime Analysis of Self-adaptive EAs

Per Kristian Lehre (University of Birmingham, GB)

We will present ongoing work on runtime analysis of self-adaptive evolutionary algorithms.

3.16 Dynamic Linear Functions

Johannes Lengler (ETH Zürich, CH)

I believe that this type of test functions are interesting, fruitful, and tractable. I will explain what they are, why I believe that it's worth studying them, and what the background is. I would like to invite others to study them, either as collaboration or independently. After the talk, there will also be a breakout session on the topic.

Related papers are:

- https://doi.org/10.1109/SSCI.2018.8628785
- https://ieeexplore.ieee.org/abstract/document/8715464
- https://doi.org/10.1145/3299904.3340309

3.17 Automated Algorithm Configuration and Selection for Theoreticians

Manuel López-Ibánez (University of Manchester, GB)

High-performing optimizers have many parameters that need to be configured. There are many benefits of automatically configuring these parameters. The problem of automatic algorithm configuration (AC) is described in a formal mathematical manner, together with a brief description of irace, which is one method for tackling it. In addition, the problem of automatic algorithm selection (AS) is described and connected to automatic algorithm configuration. Finally, topics within AC/AS of potential interest for theoreticians are highlighted with links to recent works in this direction.

3.18 What's hot in EA theory II

Pietro S. Oliveto (University of Sheffield, GB)

In this introductory talk we provided an overview of recent trends, techniques and challenges in the theoretical runtime analysis of bio-inspired optimisation heuristics. We covered the latest results and open problems concerning generational and steady-state genetic algorithms and artificial immune systems.

3.19 Dynamical Search: a Short Introduction

Luc Pronzato (Laboratoire I3S – Sophia Antipolis, FR)

Joint work of Luc Pronzato, Henry P. Wynn, Anatoly A. Zhigljavsky
Main reference Luc Pronzato, Henry P. Wynn, Anatoly A. Zhigljavsky: "Dynamical Search – Applications of Dynamical Systems in Search and Optimization: Interdisciplinary Statistics", CRC Press, 1999.

Many algorithms that aim to determine the location of a target in $\mathbb{R}^d$ (typically, the minimizer of a given function) construct a sequence of regions, of decreasing sizes, that converge towards the (fixed) target. By renormalizing the region obtained at each iteration into a fixed base region, we obtain a new representation with a fixed region and a moving target inside. It is the evolution of this moving target over iterations that defines the dynamical system, whose behavior informs us about the performance of the algorithm. All the machinery of dynamical systems can be used, including ergodic theory, and Lyapunov exponents and entropies generated (Shannon and Rényi) can be associated with measures of performance. It may happen that worst-cases have zero ergodic measure, which opens the way to an acceleration of algorithms considered as worst-case optimal. The talk is based on the book [Pronzato, L., Wynn, H., Zhigljavsky, A., 2000. Dynamical Search. Chapman & Hall/CRC, Boca Raton] and mainly focuses on line-search algorithms (like the Golden-Section method); the more difficult cases cases of the ellipsoid and steepest-descent algorithms are briefly considered.

3.20 Open problems relating to Noisy OneMax

Jonathan E. Rowe (University of Birmingham, GB)

Noisy OneMax is an interesting test problem. We know some results on the runtime for various algorithms. There are still lots of interesting open problems, and the most efficient algorithm is still unknown.

3.21 Work at the Alan Turing Institute on "The Data Science Revolution in Scientific Research"

Jonathan E. Rowe (University of Birmingham, GB)

The use of big data methods in science has curious roots, from bioinformatics and paranormal psychology, to particle physics and social economics. These methods took a strange detour via advertising, social media and playing Go, but are now finding applications in research across the breadth of science and the humanities. We will look at a range of projects where AI is transforming research practice, and the role the Alan Turing Institute is playing in this revolution. We will then consider a number of challenges this approach presents, in which some traditional philosophical questions gain unexpected practical applications.

3.22 Runtime in Integer Space Under Multiple Objectives

Günter Rudolph (TU Dortmund, DE)

The talk described a research idea.

3.23 Analysis of Evolution Strategies Applied to a More General Conically Constrained Problem

Patrick Spettel (Fachhochschule Vorarlberg – Dornbirn, AT)

Joint work of Patrick Spettel, Hans-Georg Beyer

Theoretical predictions for the behavior of evolution strategies applied to a linear objective function with a specific conical constraint have recently been derived (work presented at GECCO 2019 and FOGA 2019, among others). The specialty of that problem is that the objective function gradient's direction coincides with the cone axis. Ongoing work tries to predict the behavior of a more general conically constrained problem, in which the objective function gradient does not coincide with the cone axis. The talk presents first steps and initial results.

3.24 Runtime Analysis of Probabilistic Crowding – Results beyond OneMax

Dirk Sudholt (University of Sheffield, GB)

Joint work of Edgar Covantes Osuna, Dirk Sudholt
Main reference Edgar C. Osuna, Dirk Sudholt: "Runtime Analysis of Crowding Mechanisms for Multimodal Optimisation", IEEE Transactions on Evolutionary Computation, pp. 1–1, 2019.
URL https://doi.org/10.1109/TEVC.2019.2914606

Premature convergence is a major challenge in evolutionary computation and many diversity-preserving mechanisms have been proposed to address this. In Probabilistic Crowding, an offspring competes against its parent in a fitness-proportional selection. I showed that a (μ+1) EA with probabilistic crowding does not perform much better than random search on OneMax. We then extended our results to a much more general problem class by introducing a notation of (α, β)-bounded gradients: the gradient towards the optimum is bounded by α in a Hamming ball of radius β around global optima. The improved results show that the algorithm is unable to evolve solutions close to global optima for all functions with bounded gradients and up to exponentially many optima.

3.25 On the Linkage Equilibria of Weakly-Selective Steady-State GAs

Andrew M. Sutton (University of Minnesota – Duluth, US)

I will present some recent work (with Carsten Witt) on the time it takes a steady-state genetic algorithm using uniform crossover and weak probabilistic tournament selection to approach linkage equilibrium, i.e., a state in which sampling a string from the population is very close to the factor distribution over allele frequencies. In this state, sampling from the population is similar to drawing a sample from an EDA, which is attractive from an analysis point of view. This comes at a cost, however, as the selection is likely unreasonably weak from an optimization perspective.

3.26 Single-Iteration Evolutionary Computation (aka Fully Parallel Derivative-Free Optimization)

Olivier Teytaud (Facebook – Paris, FR)

We have a few preliminary results, and we failed to derive a good parametrization for the best performing method.

3.27 Optimal Mixing Evolutionary Algorithms

Dirk Thierens (Utrecht University, NL)

This talk discussed the GOMEA algorithm, specifically the linkage tree model.

3.28 Analysis of Artificial Genetic Representations with Neutrality

Vida Vukašinović (Jozef Stefan Institut – Ljubljana, SI)

Joint work of Vida Vukašinović, Carlos M. Fonseca, Nino Bašić

Kimura's theory of evolution suggests the possibility of occurrence of so called neutral networks. The potential of neutral networks to establish alternative paths for the evolution of a population, and to lead to improved search quality, is the main motivation for the use of redundant representations in evolutionary computation, although not all redundant representations exhibit neutrality. We prepared a solid mathematical formalization of the binary representations developed by Fonseca and Correia (2005) and their equivalence classes. Those representations can exhibit various degrees of neutrality, connectivity, locality, and synonymity, all of which are properties known (or believed) to influence the performance of evolutionary algorithms. Based on this, we developed an efficient algorithm allowing for the exhaustive enumeration of a family of 15-bit representations involving 4 bits of redundancy. The questions of how to identify good, or at least promising, representations in such a large database, and how to automate their (theoretical and practical) evaluation, remain open.

3.29 What's Hot In EA Theory I

Carsten Witt (Technical University of Denmark – Lyngby, DK)

The purpose of this introductory talk was to give an overview of recent trends, techniques and challenges in the theoretical runtime analysis of evolutionary algorithms (EAs). We covered the exact analysis of EAs via drift theory, estimation-of-distribution algorithms and self-adjusting EAs. Topics for future research included monotone functions, multivariate estimation-of-distribution algorithms and a more comprehensive theory of self-adjusting EAs.

3.30 Stochastic global optimization (SGO)

Anatoly Zhigljavsky (Cardiff University, GB)

Main reference Anatoly A. Zhigljavsky: "Theory of Global Random Search". Kluwer Academic Press, Dordrecht, 1991, pp xviii+342
URL https://doi.org/10.1007/978-94-011-3436-1
Main reference Anatoly A. Zhigljavsky, Antanas Zilinskas: "Stochastic Global Optimization", Springer, 2008.

The talk was devoted to some open issues in the theory of global random search, in particular, to the rate of convergence of global random search algorithms in large dimensions and to the theory of evolutionary global random search algorithms. In particular, it was shown that for a large class of stationary evolutionary algorithms the asymptotic distribution of points approaches a stationary limiting distribution, which generalizes the celebrated Gibbs distribution.

4 Working groups

4.1 Breakout Session: Benchmarking – Best Practices and Open Issues

Thomas Bartz-Beielstein (TH Köln, DE)

The goal of this breakout session was to coordinate the working group for writing a commented survey article on "Benchmarking – Best Practices and Open Issues". The aim is to have a survey that is broadly accepted in the community. The outcome of the breakout session was a plan for collecting information and writing the article with the aim of publishing it in 2020.

4.2 Breakout Session: Multiobjective Optimization

Dimo Brockhoff (INRIA Saclay – Palaiseau, FR), Benjamin Doerr (Ecole Polytechnique – Palaiseau, FR), Carola Doerr (Sorbonne University – Paris, FR), Manuel López-Ibánez (University of Manchester, GB), Rolf H. Möhring (TU Berlin, DE), Günter Rudolph (TU Dortmund, DE), Dirk Thierens (Utrecht University, NL), Markus Wagner (University of Adelaide, AU), and Elizabeth Wanner (Aston University – Birmingham, GB)

Time and date: 22.10.2019, 14:30 – 15:30

Participants: Dimo Brockhoff, Benjamin Doerr, Carola Doerr, Manuel López-Ibánez, Rolf H. Möhring, Günter Rudolph, Dirk Thierens, Markus Wagner, Elizabeth Wanner

We started with a short round, in which every participant briefly stated their involvement with multiobjective optimization and theoretical analyses in this context in particular. It turned out that the nine participants have quite heterogeneous backgrounds: about 2/3 claimed a theoretical background (about 1/3 did not do any theory before) and about 1/3 claimed experiences with non-CI methods.

The discussed topics in this working group can be categorized into previously researched (theoretical multiobjective optimization) topics and potential topics for future research. We will detail them in the following subsections but can conclude already here that, before starting to analyze any algorithm (runtime), we have to understand the underlying fundamental problems first.

4.2.1 Previous Research Topics

Compared to single-objective optimization, the theory of (population-based) multiobjective optimization is still in its infancy. Within the short time of the breakout session, we identified only the following, non-exhaustive list of topics that have been touched by previous research:

- fundamental aspects (not related to an algorithm)
  - approximation guarantees
  - optimal p-distributions
  - subset selection
  - properties of quality indicators
  - In discrete problems, the approximation of the Pareto front is polynomially equivalent to the approximation of the ideal point. This has been shown in [1] and means that – at least in theory – algorithms might just concentrate on computing or approximating the ideal point and then use the polynomial transformations of this paper to obtain an approximation to the whole Pareto front.
- computational geometry related problems such as hypervolume computations, see https://simco.gforge.inria.fr/doku.php?id=openproblems for a detailed list
- first runtime analyses

4.2.2 Topics for Future Research

In the remaining time of the breakout session, we collected potential topics for future research such as

- some topics where we don't know the complexity
  - computation of the hypervolume indicator in high dimension
  - algorithms for (bounded) archiving
- similarities between single- and multiobjective optimization, for example, what can be learned/transferred between the two scenarios? What are the differences?
- tradeoffs between different algorithm types (when is which better, e.g. correlation between the objectives); example: (Pareto) local search vs. scalarization, which one is better (and when)?
- Pareto-compliant indicators: how much can they disagree?
- already understanding properties of objective functions is hard in the multiobjective case:
  - The multiobjective quadratic assignment problem has very different instances but the instances of the (random) multiobjective knapsack problem are much less different. The question is why?
  - How do landscapes look like for certain quality indicators and/or different operators? In this context, Manuel brought up the study on NK landscapes in which it was shown empirically that more local optima exist with respect to dominance (between sets) compared to the number of local optima if the hypervolume indicator is the set quality criterion. Also the epsilon indicator shows more local optima compared to the hypervolume indicator. See [2] for details.

References

1 C. Büsing, K.-S. Goetzmann, J. Matuschke, and S. Stiller. Reference points and approximation algorithms in multicriteria discrete optimization. *European Journal of Operational Research*, 260(3):829–840, 2017.

2 A. Liefooghe, M. López-Ibáñez, L. Paquete, and S. Verel. Dominance, epsilon, and hypervolume local optimal sets in multi-objective optimization, and how to tell the difference. In H. E. Aguirre and K. Takadama, editors, *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO 2018*, pages 324–331. ACM Press, New York, NY, 2018. 10.1145/3205455.3205572.

4.3 Breakout Session: Mixed-Integer-Nominal Optimization

Thomas Bäck (Leiden University, NL)

This breakout session discussed and collected ideas for *mixed-integer-nominal optimization*. Different aspects have been discussed. As a starting point, it is interesting to look into the state of the art of mixed-integer nonlinear programming (MINLP). An idea for handling such problems in evolution strategies is to have a joint covariance matrix between integer and continuous variables. A further interesting question is what the typical problems in this area are. Different such problems were collected: There is work concerning landscape features [1]. Further problems were mentioned including optical multilayer systems (thickness and materials), car body safety optimization (thickness and materials), test case generation in software engineering (arguments for functions, Evosuite). For the theory community, a "standard" problem in this domain could be interesting.

References

1 Carola Doerr, Johann Dréo, and Pascal Kerschke. Making a case for (hyper-)parameter tuning as benchmark problems. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion (GECCO'19)*, pages 1755–1764. ACM, 2019. 10.1145/3319619.3326857. URL https://doi.org/10.1145/3319619.3326857.

4.4 Breakout Session: Open Problems

Benjamin Doerr (Ecole Polytechnique – Palaiseau, FR) and Frank Neumann (University of Adelaide, AU)

The open problem session attracted a good 20 participants and nine open problems from all subdisciplines of the theory of randomized search heuristics. Each problem was presented in at most five minutes and then discussed for as long as the participants had to say something. A number of contradicting conjectures were made, which promises that we will soon see some interesting progress in one direction or the other. The problems can be found on the seminar page.

4.5 Breakout Session: Analysis of Artificial Genetic Representations with Neutrality

Carlos M. Fonseca (University of Coimbra, PT) and Vida Vukašinović (Jozef Stefan Institut – Ljubljana, SI)

The potential of neutral networks to establish alternative paths for the evolution of a population, and to lead to improved search quality, is the main motivation for the use of redundant representations in evolutionary computation, although not all redundant representations exhibit neutrality. We prepared a solid mathematical formalization of the binary representations developed by Fonseca and Correia (2005) and their equivalence classes. Those representations can exhibit various degrees of neutrality, connectivity, locality, and synonymity, all of which are properties known (or believed) to influence the performance of evolutionary algorithms. Based on this, we developed an efficient algorithm allowing for the exhaustive enumeration of a family of 15-bit representations involving 4 bits of redundancy. The representation database obtained in this manner contains over 4.58×10^{10} canonical representatives, each of which representing up to 20160 different representations. The questions of how to identify good, or at least promising, representations in such a large database, and how to automate their (theoretical and practical) evaluation, remain open. The aim of the proposed breakout session was to:

- Discuss current hypotheses about the role of neutrality in evolutionary search and how they may be investigated using this data
- Explore collaborations related to the runtime analysis of evolutionary algorithms based on such representations on simple problems
- Discuss other research opportunities offered by the availability of a database of this kind.

In the beginning, we explained the proposed representations into adequate details. Carlos showed the database as well he presented current feasibilities and obstacles in the database manipulation. During discussion on runtime analysis a justification of runtime analysis of evolutionary algorithms on simple problems based on the proposed representations was put under the question. Main arguments were that by using such representations for simple problems we could not expect better algorithm performance and proving that (1+1)-EA needs $\Omega(n \log n)$ function evaluations is not an impressive result. Nevertheless, we agreed that for deeper understanding what is the influence of proposed representations on the algorithm performance such study is essential. We were able to identify some concrete representations for which runtime analysis of (1+1)-EA on OneMax seems doable. By this an opportunity for further collaboration related to the runtime analysis with the researchers present at the breakout session is opened.

4.6 Breakout Session: Connection Between ES / EDA Analysis and Stochastic Approximation / ODE Theory

Tobias Glasmachers (Ruhr-Universität Bochum, DE)

Stochastic approximation and ODE methods are powerful tools to analyze stochastic algorithms that are formalized as a stochastic approximation of the solution of an underlying Ordinary Differential Equation. In this session we want to discuss how algorithms like Evolution Strategies (ES) (and at least some EDAs) can be casted in the framework of stochastic approximation methods and whether standard ODE methods apply or what is missing in current ODE method to be able to apply it to analyze ES and EDAs.

4.7 Breakout Session: Measuring Optimization Progress in an Invariant Way for Comparison-Based Algorithms

Tobias Glasmachers (Ruhr-Universität Bochum, DE)

Comparison-based or value-free algorithms ignore actual objective function values and instead only use pairwise "better or worse" comparisons. This property renders them invariant to strictly monotonically increasing transformations of objective values. Therefore, measuring quality and optimization progress in terms of function values is inappropriate since it does not exhibit the same invariance properties – possibly, unless there is a clear meaning attributed to these values in an application. An alternative approach is to consider the distance to the optimum instead. That choice is equally problematic unless the (then trivial) objective function is itself a function of that distance. The currently most promising way out of this dilemma is to consider the size (continuous case: Lebesgue measure) of sub-level sets.

4.8 Breakout Session: Invariance

Nikolaus Hansen (INRIA Saclay – Palaiseau, FR) and Anne Auger (INRIA Saclay – Palaiseau, FR)

The invariance breakout gathered participants from the different domains of research present at the seminar. The role of invariance and its importance was acknowledged for algorithm design. In the discrete search domain relatively mild invariance assumptions have lead to a proof of lower runtime bounds. The question was raised whether more results of that type should be expected or attempted. Invariance has also been instrumental for convergence proofs on continuous search spaces via Markov chains. The specific questions on how to model randomness was raised, where methods from stochastic approximation may prove to be useful in convergence proofs for evolutionary algorithms. We also scrutinized specific formulations of invariance, whether to consider the algorithm state in the formalization, and whether the (weak) notion of asymptotic invariance could be helpful.

4.9 Breakout Session: Drift Theory

Martin S. Krejca (Hasso Plattner Institute – Potsdam, DE)

We started with brainstorming some settings in which new drift theorems would be helpful but do not exist yet. The first proposal was a fixed-budget scenario, where one is interested in bounding the expected progress a process has made after a certain (known) time. The discussion suggested that tail bounds on the probability of the process not having reached a target state yet play an important role.

The next setting was drift in multiple dimensions, which is interesting when considering, for example, multiple (independent) processes that should all hit a target value. It was mentioned that the idea of super- and submartingales also generalizes to vectors of random variables. This might yield a useful approach to tackle this problem.

The discussion then moved to a scenario that frequently occurs in multiplicative drift: while the process is far away from the target, its drift is the dominating factor for the expected run time. However, when getting close to the target, the variance of the process is more crucial. It was discussed how the analyses of these two regimes could be combined. The conclusion was to consider the expected return time of the process in the regime of dominating variance, which basically amounts to a restart argument.

Afterward, a drift-like setting for stochastic domination was discussed. The idea was to consider a process where not the expected difference within a single time step is bounded (like in classical drift) but instead a stochastic domination is observed. It was decided that the problem needs to be formalized more rigorously in order to make precise statements.

In the end, the initial idea of drift in the fixed-budget setting was considered in a bit more detail. However, no final conclusion was reached until the session ended.

4.10 Breakout Session: Passing It On

Timo Kötzing (Hasso Plattner Institute – Potsdam, DE)

URL https://github.com/TeachingMetaheuristics

From time to time we all get PhD students or Master Students interested in joining our kind of research. Sometimes we want to teach our subject as a lecture. And also, sometimes researchers from other areas would like to understand better what we do. There are several books available to help such projects along, as well as plenty of other material: talks, scripts, collections of exercises and so on.

In this breakout we discussed in what fashion we could make all these resources widely available for anyone to use. We decided on the following:

- Open a dedicated GitHub repository.
- Let anyone submit more material to this repository.
- The material should be tagged and/or uploaded with a certain structure for easy browsing.
- Anybody who wants can write a “guide” which leads through a subset of the material, aiming at providing a certain expertise. For example
  - “Guide to doing run time analysis of search heuristics” could point to material covering arithmetic and stochastic inequalities, followed by pointing to drift theorems and some easy sample applications.

- Naturally, different material will have a different angle.

What we imagine this resource could be:
- An introduction to the field if read start to finish.
- Reading assignments for classes on the topic.
- A collection of useful assignments.
- A resource for looking up central results.
- Highly modular.
- Multi-authored, yet curated.

We agreed on the following.
- Timo and Thomas W. set up a GitHub.
- Timo sends an Email to all Dagstuhl seminar participants, inviting them to add their materials.
- Once some material is there, Timo sends an email inviting guides.

The github repository can be found at https://github.com/TeachingMetaheuristics.

4.10.1 Addendum by T. Weise

Maybe interesting in this context might be an automated tool chain for writing electronic books. If you host the book's sources in GitHub, the book can get automatically compiled and published to pdf, html, azw3, and epub and uploaded upon each commit: slides, early stage/incomplete example book project.

4.10.2 Addendum by M. Buzdalov

I have been teaching evolutionary computation for maybe three years already, and my other courses are about algorithms and data structures. This inevitably produced a crossover between them, in particular, in a form of *automatically tested programming assignments* for various aspects of evolutionary computation. These should be designed in such a way that solving them using non-evolutionary methods shall be inefficient or impossible. Surprisingly, *there are some*, especially with gray-box techniques. I think that:
- having more such problems is beneficial both for teaching the subject and for ourselves being confident that what we do is interesting for the general public;
- this would be a good practical driver to develop better algorithms and teach the subject properly.

4.11 Breakout Session: The Purpose of Theory Research

Timo Kötzing (Hasso Plattner Institute – Potsdam, DE)

We all do our research, and we all have a good gut feel for what constitutes good research. In this breakout, we brought this gut feel a bit more into the conscious realm and discuss what makes a result "good" and what would be considered less interesting. For example, knowing the lead constant of the expected optimization time of the 1+1 EA on OneMax is (to me?!) not so much of interest in itself; rather, (i) we gain understanding of the inner working principles of the 1+1 EA which (ii) allows us to get a *feel* for many other problems as well, the result generalizes, (iii) let's us dig deeper in related areas after having understood this part. It also (iv) lead us to develop tools (such as drift theory) which are applicable in other contexts as well.

In this breakout we determined the following ingredients for a paper to be worthwhile research.

- **Proper Execution**: The paper is well-written, experiments are clear, ideas are given, proofs are rigorous.
- **Connection to Scientific Community**: Works on topics also others are interested in, discusses own research in context of others', stimulates further research.
- **Academic Honesty**: Explains limits of applicability, does not oversell, does not cheat.
- **Novelty**: The work contains a new idea or a new angle, possibly disproofs commonly held beliefs.
- **Motivation**: Two core sources of a good motivation can be **applicability** (either internally for further research or externally providing valuable input to other scientific communities) and **explanation** of phenomena (also and importantly from other sciences); further worthy kinds of motivation include giving the broader picture, unifying results, settling important questions.

4.12 Breakout Session: Competitive Co-evolution

Per Kristian Lehre (University of Birmingham, GB)

Since long, the field of evolutionary computation has demonstrated empirically that competitive co-evolutionary algorithms can provide state-of-the art solution to certain types of optimisation problems, such as design of sorting networks. However, co-evolutionary algorithms often show pathological behaviour, such as disengagement, loss of gradient, cycling, and overspecialisation. There is currently no theory able to predict and explain this behaviour.

This breakout session explored the potential for runtime analysis of competitive co-evolutionary algorithms. To clarify what runtime means in this context, it is necessary to specify a solution concept, a class of co-evolutionary algorithms, and a class of games.

One well-known solution concept is **Nash equilibrium**. We also discussed **maximin-optimisation**, where we consider interactions between a set of candidate solutions $\mathcal{X} = \{0,1\}^n$ and a set of tests $\mathcal{Y} = \{0,1\}^n$, defined by an interaction function

$$g : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}.$$

Here, $g(x, y)$ gives the *performance* of solution $x \in \mathcal{X}$ on test $y \in \mathcal{Y}$. Our goal is to find a solution $x \in \mathcal{X}$ which maximises the function

$$h(x) := \min_{y \in \mathcal{Y}} g(x, y), \tag{1}$$

i.e., to find the solution x which performs best when evaluated with respect to its worst-case test y.

Jon Rowe suggested a co-evolutionary setting with an underlying pseudo-Boolean function $f : \{0,1\}^n \rightarrow \mathbb{R}$ and two competing (1+1) EAs. Algorithm A attempts to optimise the function f by finding a good search point x, while Algorithm B attempts to fool Algorithm A by choosing a "wildcard" y. The algorithms interact via a function g defined as below, which Algorithm A attempts to maximise, and Algorithm B attempts to minimise:

$$g(x, y) = f\left(\max(x_1, y_1), \ldots, \max(x_n, y_n)\right). \tag{2}$$

Algorithm 1 Co-evolutionary (1+1) EA

Require: Fitness function $f : \{0,1\}^n \to \mathbb{R}$
Require: Interaction function $g : \{0,1\}^n \times \{0,1\}^n \to \mathbb{R}$ as in Eq. (2)
Sample initial search points $x, y \sim Unif(\{0,1\}^n)$
while stopping condition not met **do**
  Obtain x' by flipping each bit in x with prob $1/n$.
  Obtain y' by flipping each bit in y with prob $1/n$.
  if $g(x', y') \geq g(x, y)$ **then**
    $x \leftarrow x'$
  end if
  if $g(x', y') \leq g(x, y)$ **then**
    $y \leftarrow y'$
  end if
end while

It is an open problem to determine for what functions f Algorithm 1 finds an optimal solution x in expected polynomial time. It might be necessary to choose strict inequalities when updating the current search points x and y.

Per Kristian Lehre suggested a framework for population-based co-evolutionary algorithms (Algorithm 2), where a specific algorithm is obtained by choosing a specific operator $\mathcal{D}$. Implicitly, the dynamics is governed by a two-player normal form game with payoff matrices G and H respectively. We make the following assumptions:

1. The strategy spaces $\mathcal{X}$ and $\mathcal{Y}$ are finite and exponentially large in some parameter n, e.g., $\mathcal{X} = \mathcal{Y} = \{0,1\}^n$. Thus, G and H are non-differentiable.
2. The functions G and H can be non-convex.
3. The algorithm is limited to "black box access" to G and H.
4. We have a "solution concept" given by a subset $\mathcal{S} \subset \mathcal{X} \times \mathcal{Y}$
5. $\mathcal{D}$ is non-deterministic, i.e., we need to take into account stochastic effects.

Algorithm 2 Co-evolutionary algorithm

Require: Payoff matrices $G, H : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$
Require: Population size $\lambda \in \mathbb{N}$
for i in $1, \ldots, \lambda$ **do**
  $P_0(i) := (x, y, G(x, y), H(x, y))$ where $(x, y) \sim Unif(\mathcal{X} \times \mathcal{Y})$.
end for
for t in $0, \ldots, \lambda$ **do**
  for i in $1, \ldots, \lambda$ **do**
    $P_{t+1}(i) := (x, y, G(x, y), H(x, y))$ where $(x, y) \sim \mathcal{D}(P_t)$
  end for
end for

First, in steps 1–3, the algorithm samples λ initial pairs of learners and teachers (x, y) uniformly at random, and evaluates the payoff $G(x, y)$ of the learner x, and the payoff of the teacher $H(x, y)$. In each generation t, in steps 5–7, the algorithm samples and evaluates λ new pairs of learners and teachers from a probability distribution $\mathcal{D}(P_t)$ which depends on the current interaction outcomes $P_t \in (\mathcal{X} \times \mathcal{Y} \times \mathbb{R} \times \mathbb{R})^\lambda$.

A (pure) solution concept corresponds to a subset $\mathcal{S} \subseteq \mathcal{X} \times \mathcal{Y}$ of the strategy space. The objective of the algorithm is to discover a pair of individuals (x, y) in this set.

▶ **Definition 1** (Runtime).

$$T_{A,\mathcal{S}} := \min\{t \in \mathbb{N} \mid \exists j \in [\lambda] \text{ such that } P_t(j) \in \mathcal{S}\}.$$

Per Kristian also suggested the following maximin-optimisation benchmark problem. The utility function for the prey is $u_1(x, y) = d(x, y)$, while the utility function for the predator is $u_2(x, y) = -d(x, y)$, where for any parameter $\varepsilon \geq 0$,

$$d(x, y) := (|y|_1 - \varepsilon|x|_1)^2 ,$$

and $|z|_1 := \sum_{i=1}^n z_i$ for all bitstrings $z \in \{0, 1\}^n$.

There is a large literature on related concepts in biology, economics, and theoretical computer science. Vivek S. Borkar discussed results from evolutionary game theory.

4.13 Breakout Session: Dynamic Linear Functions

Johannes Lengler (ETH Zürich, CH)

Dynamic Linear Functions are a benchmark where we use a linear function with positive weights, but every few rounds (or every round), the weights are redrawn and everything in the population is re-evaluated. This sounds rather trivial, but it is not. I believe that this is an extremely rich and rewarding topic, and in the talk preceding the breakout session, I will explain where this belief comes from. In a nutshell, they are the easier (less technical) siblings of the monotone functions that have surprised us so many times in the least years.

Goal There are three goals that I hope to achieve:
- Discuss which settings of these functions are most interesting. (E.g., how often should the weights change.)
- Discuss which research questions would be most interesting.
- Motivate other people to work on the topic, either in a collaboration with me, or independently.

Length 30-60 minutes.

Method Group discussion.

Outcome We discussed first how we could categorize the algorithms that fail on dynamic linear functions, or monotone functions. We also discussed the black-box complexity of dynamic linear functions, which is in $\Omega(n/\log n) \cap O(n)$.
We then moved on to discuss algorithms that would be interesting to study on dynamic linear functions. A recurring theme were EDAs like the compact Genetic Algorithm cGA. There were split opinions on how efficient the cGA would be on these benchmarks. Of

particular interest might be whether dynamic linear functions could actually be easier than the static instance of BinVal, i.e., whether noise might actually make optimization easier in this case.

John Rowe pointed out that dynamic linear function have a very peculiar aspect in that the noise level increases as the algorithm moves closer to the optimum, since one-bits contribute noise, whereas zero-bits don't. Finally, we discussed possible extensions, in particular a model in which not all weights are redrawn every round, but rather only a subset of weights is redrawn.

4.14 Breakout Session: Algorithm Configuration and Selection

Pietro S. Oliveto (University of Sheffield, GB)

Main reference George T. Hall, Pietro Simone Oliveto, Dirk Sudholt: "On the impact of the cutoff time on the performance of algorithm configurators", in Proc. of the Genetic and Evolutionary Computation Conference, GECCO 2019, Prague, Czech Republic, July 13-17, 2019, pp. 907–915, 2019.
URL http://dx.doi.org/10.1145/3321707.3321879

In this brief breakout session we discussed the state-of-the-art in the time complexity analysis of algorithm configurators. Most of the discussion concerned which performance measures should be preferred to compare the effectiveness of different parameter settings. Time to optimality, Best identified fitness, and time to fixed targets were considered.

4.15 Breakout Session: Competitions and Benchmarking

Olivier Teytaud (Facebook – Paris, FR) and Carola Doerr (Sorbonne University – Paris, FR)

GECCO and other conferences are hosting several workshops on benchmarking evolutionary algorithms. In addition, a number of competitions are proposed. At the moment, there is little coordination between the workshops and competitions, and we have discussed if it makes sense to coordinate efforts and/or to share best practices and pitfalls.

In the first session the focus of the discussion has centered around the question whether competitions are useful for our understanding of algorithmic behavior, or whether they encourage too much overfitting. In the discussion, most/all participants agreed that contributions to the benchmark environment (e.g., suggestion of new benchmark problems, additional features for a software-based analysis, a critical discussion of different statistics, etc.) are at least as important as the development of high-performing algorithms. Participants agree that such ideas should be "rewarded" as well. A comparison has been made to the Pytorch, which is used in Machine Learning, and which benefits from a user-friendly platform. Another examples that has been mentioned in this context is OpenML, which has similar goals than what we consider a widely accepted benchmarking environment.

At the end of the session we have discussed the idea to have an award committee, which is different from and independent of the organizing committee of the competition, and which judges the contributions made to EC-centered benchmarking environments.

4.16 Breakout Session: One-Shot Optimization

Olivier Teytaud (Facebook – Paris, FR) and Carola Doerr (Sorbonne University – Paris, FR)

In one-shot optimization, aka *single-iteration evolution* or *fully parallel optimization*, the user selects a population, evaluates it, and has to base all future decisions only on the quality of these points. In recent work, O. Teytaud and co-authors have analyzed the setting in which an optimal solution is chosen at random from a Gaussian distribution. They could prove that, unlike one might have guessed, it is better to sample only one (namely, the center of the distribution) rather than sampling n times from the same Gaussian distribution [1]. In the breakout session we have proven that sampling the middle point is not optimal. We have started to compute the optimal distribution, but will need to resume this discussion offline.

Participants: Thomas Bartz-Beielstein, Alexandre Chotard, Carola Doerr, and Olivier Teytaud.

References

1 Marie-Liesse Cauwet, Camille Couprie, Julien Dehos, Pauline Luc, Jérémy Rapin, Morgane Rivière, Fabien Teytaud, and Olivier Teytaud. Fully parallel hyperparameter search: Reshaped space-filling. *CoRR*, abs/1910.08406, 2019. URL http://arxiv.org/abs/1910.08406.

4.17 Breakout Session: Permutation-based Problems

Christine Zarges (Aberystwyth University, GB)

The aim of this breakout session was to discuss current work and potential research directions for permutation-based problems.

Starting with a brief recap of the GECCO 2019 workshop on "Evolutionary Computation for Permutation Problems", the discussion first evolved around different types of permutation-based problems (e.g., total ordering, partial ordering, adjacency) and example problems (e.g., travelling salesperson, linear ordering, linear and quadratic assignment, OneMax variants as introduced in the talk by Maxim Buzdalov, Lyndon factorisation).

Afterwards the group focussed its discussion on metrics and permutation spaces based on the following two publications:

- Ekhine Irurozki: Sampling and learning distance-based probability models for permutation spaces. PhD Thesis, University of the Basque Country, 2014.
- Tommaso Schiavinotto and Thomas Stützle: A review of metrics on permutations for search landscape analysis. Computers & OR 34(10): 3143-3153 (2007)

Carlos Fonseca also pointed out that the API developed by working group 4 of COST Action CA15140 contains some common neighbourhood definitions that could be useful for future work.

Participants: Francisco Chicano, Anton V. Eremeev, Carlos Fonseca, Andrei Lissovoi, Dirk Thierens, Christine Zarges

5 Schedule

Monday

– 09:00	Breakfast
09:00 – 09:15	Welcome and seminar opening
09:15 – 10:00	Participant introduction I
10:00 – 10:30	Coffee break
10:30 – 11:00	**Carsten Witt** on *What's Hot in EA Theory I*
11:00 – 11:20	**Benjamin Doerr** on *Genetic Drift in EDAs*
11:20 – 12:00	Participant introduction II
12:15 – 13:30	Lunch
13:30 – 14:30	Time for individual discussions
14:30 – 15:00	**Luc Pronzato** on *Dynamical Search*
15:00 – 15:20	Participant introduction III
15:20 – 15:30	Short announcement concerning the group work
15:30 – 16:00	Cake
16:00 – 16:15	Organization of group work
16:15 – 17:15	**Breakout Sessions:**
	Passing It On (Timo Kötzing)
	Measuring Optimization Progress in an Invariant Way for Comparison-Based Algorithms (Tobias Glasmachers)
17:15 – 17:45	Participant introduction IV
17:45 – 18:00	Debrief from the breakout sessions
18:00 – 19:00	Dinner
19:30 –	Opening of the art exhibit "Lost Places" by the German artist Winfried Groke

Tuesday

– 09:00	Breakfast
09:00 – 09:20	**Johannes Lengler** on *Dynamic linear functions*
09:20 – 09:40	**Vida Vukašinović** on *Analysis of artificial genetic representations with neutrality*
09:40 – 10:00	**Olivier Teytaud** on *Single-iteration evolutionary computation, also known as fully parallel derivative-free optimization*
10:00 – 10:30	Coffee break
10:30 – 11:00	**Pietro Oliveto** on *What's hot in EA theory II*
11:00 – 11:30	**Vivek Borkar** on *Overview of stochastic approximation and related schemes*
11:30 – 12:00	**Youhei Akimoto** on *Expected runtime bounds for* $(1+1)$*-ES*
12:15 – 13:30	Lunch
13:30 – 14:00	Time for individual discussions
14:00 – 14:30	Organization of group work
14:30 – 15:30	**Breakout Sessions:**
	Neutral Representation (Carlos Fonseca, Vida Vukašinović)
	IGO/Stochastic Optimization (Anne Auger, Tobias Glasmachers)
	Multi-Objective Optimization (Dimo Brockhoff)
	Dynamic Linear Functions (Johannes Lengler)
15:30 – 16:00	Cake
16:00 – 16:10	Flash talk: **Günter Rudolph** on *Runtime in integer space under multiple objectives*
16:10 – 17:30	**Breakout Sessions:**
	Drift Analysis (Martin Krejca)
	Benchmarking and Competition (Olivier Teytaud, Carola Doerr)
17:30 – 18:00	Debrief and announcements
18:00 – 19:00	Dinner
19:00 –	Individual discussions

Wednesday

– 09:00	Breakfast
09:00 – 09:20	**Patrick Spettel** on *Analysis of evolution strategies applied to a more general conically constrained problem*
09:20 – 09:40	**Anne Auger** on *A unified invariance formalism for discrete and continuous optimization*
09:40 – 10:00	**Niko Hansen** on *A (general) definition of invariance*
10:00 – 10:30	Coffee break
10:30 – 10:50	**Anatoly Zhigljavsky** on *Stochastic global optimization (SGO)*
10:50 – 11:10	**Martin Krejca** on *The UMDA on LeadingOnes revisited*
11:10 – 11:30	**Dirk Sudholt** on *Runtime analysis of diversity mechanisms – recent results*
11:30 – 12:00	Organization of group work
12:15 – 13:30	Lunch
13:30 – 15:30	Hike
15:30 – 16:00	Cake
16:00 – 18:00	**Breakout Sessions:**
	Algorithm Configuration and Selection (Pietro Oliveto)
18:00 – 19:00	Dinner
19:00 –	Individual discussions

Thursday

– 09:00	Breakfast
09:00 – 09:20	**Francisco Chicano** on *Dynastic Potential Crossover*
09:20 – 09:40	**Denis Antipov** on *Precise Analysis for Plateaus*
09:40 – 10:00	**Hans-Georg Beyer** on *Evolution Strategies are NOT Gradient Followers*
10:00 – 10:10	Group Picture
10:10 – 10:45	Coffee break
10:45 – 11:05	**Maxim Buzdalov** on *Variations on the Theme of the* $(1 + (\lambda, \lambda))$ *GA*
11:05 – 11:25	**Per Kristian Lehre** on *Runtime Analysis of Self-adaptive EAs*
11:25 – 11:45	**Anton Eremeev** on *On potential for transfer of results from theory of evolutionary algorithms to biology*
11:45 – 11:55	**Jon Rowe** on *Open Questions Relating to Noisy OneMax*
12:15 – 13:30	Lunch
13:30 – 14:30	Time for individual discussions
14:30 – 15:30	**Breakout Sessions:**
	Permutation-based problems (Christine Zarges)
	The Purpose of Theory Research (Timo Kötzing)
	Benchmarking Survey (Thomas Bartz-Beielstein)
	Invariance (Niko Hansen, Anne Auger)
15:30 – 16:00	Cake
16:00 – 17:30	**Breakout Sessions:**
	Mixed-Integer-Nominal Optimization (Thomas Bäck)
	Open Problems (Benjamin Doerr and Frank Neumann)
	One-shot Optimization (Olivier Teytaud)
	Competitive Co-evolution (Per Kristian Lehre)
17:30 – 18:00	Debrief from breakout sessions
18:00 – 19:00	Dinner
19:30 – 20:00	**Jon Rowe** on *Work at the Alan Turing Institute on "The Data Science Revolution in Scientific Research*
20:00 – 20:30	Individual discussions
20:30 –	Wine & cheese party

Friday

– 09:00	Breakfast
09:20 – 09:40	**Manuel López-Ibáñez**
	Automated Algorithm Configuration and Selection for Theoreticians
09:40 – 10:00	**Andrew Sutton** on *On the Linkage Equilibria of Weakly-Selective Steady-State GAs*
10:00 – 10:10	**Dirk Thierens** on *Optimal Mixing Evolutionary Algorithms*
10:10 – 10:30	Coffee break
10:30 – 10:40	**Arina Buzdalova** on *Challenges of mutation rate control in* $(1+\lambda)$ *EA*
10:40 – 11:00	**Niko Hansen** on *Gradient Descent and Evolution Strategies are Almost the Same*
11:00 – 11:20	**Alexandre Chotard** on *Adaptation of a Sampling Distribution for Metropolis-Hastings*
11:20 – 12:00	Closing session, feedback, and goodbye
12:15 – 13:30	Lunch
13:30 –	Individual departures

Participants

- Youhei Akimoto
University of Tsukuba, JP
- Denis Antipov
ITMO University – St. Petersburg, RU
- Anne Auger
INRIA Saclay – Palaiseau, FR
- Thomas Bäck
Leiden University, NL
- Thomas Bartz-Beielstein
TH Köln, DE
- Hans-Georg Beyer
Fachhochschule Vorarlberg – Dornbirn, AT
- Vivek Shripad Borkar
Indian Institute of Technology Bombay, IN
- Dimo Brockhoff
INRIA Saclay – Palaiseau, FR
- Maxim Buzdalov
ITMO University – St. Petersburg, RU
- Arina Buzdalova
ITMO University – St. Petersburg, RU
- Francisco Chicano
University of Málaga, ES
- Alexandre Chotard
Calais University, FR
- Benjamin Doerr
Ecole Polytechnique – Palaiseau, FR
- Carola Doerr
Sorbonne University – Paris, FR
- Anton V. Eremeev
Institute of Scientific Information for Social Sciences RAS – Moscow, RU
- Carlos M. Fonseca
University of Coimbra, PT
- Tobias Friedrich
Hasso-Plattner-Institut – Potsdam, DE
- Tobias Glasmachers
Ruhr-Universität Bochum, DE
- Nikolaus Hansen
INRIA Saclay – Palaiseau, FR
- Thomas Jansen
Aberystwyth University, GB
- Timo Kötzing
Hasso-Plattner-Institut – Potsdam, DE
- Martin S. Krejca
Hasso-Plattner-Institut – Potsdam, DE
- Per Kristian Lehre
University of Birmingham, GB
- Johannes Lengler
ETH Zürich, CH
- Andrei Lissovoi
University of Sheffield, GB
- Manuel López-Ibánez
University of Manchester, GB
- Rolf H. Möhring
TU Berlin, DE
- Frank Neumann
University of Adelaide, AU
- Pietro S. Oliveto
University of Sheffield, GB
- Luc Pronzato
Laboratoire I3S – Sophia Antipolis, FR
- Jonathan E. Rowe
University of Birmingham, GB
- Günter Rudolph
TU Dortmund, DE
- Ofer M. Shir
Tel-Hai College – Upper Galilee, IL
- Patrick Spettel
Fachhochschule Vorarlberg – Dornbirn, AT
- Dirk Sudholt
University of Sheffield, GB
- Andrew M. Sutton
University of Minnesota – Duluth, US
- Olivier Teytaud
Facebook – Paris, FR
- Dirk Thierens
Utrecht University, NL
- Vida Vukašinovic
Jozef Stefan Institut – Ljubljana, SI
- Markus Wagner
University of Adelaide, AU
- Elizabeth Wanner
Aston University – Birmingham, GB
- Thomas Weise
Hefei University, CN
- Carsten Witt
Technical University of Denmark – Lyngby, DK
- Xin Yao
Southern Univ. of Science and Technology – Shenzen, CN
- Christine Zarges
Aberystwyth University, GB
- Anatoly Zhigljavsky
Cardiff University, GB