

Theory of Randomized Optimization Heuristics

Anne Auger^{*1}, Tobias Glasmachers^{*2}, Martin S. Krejca^{*3},
Johannes Lengler^{*4}, and Alexander Jungeilges^{†5}

1 Inria Saclay Ile-de-France, Palaiseau, FR. anne.auger@inria.fr

2 Ruhr University Bochum, DE. tobias.glasmlachlers@ini.rub.de

3 Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, FR.
martin.krejca@polytechnique.edu

4 ETH Zürich, CH. johannes.lengler@inf.ethz.ch

5 Ruhr University Bochum, DE. alexander.jungeilges@ini.ruhr-uni-bochum.de

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 24271 “Theory of Randomized Optimization Heuristics”, which marks the twelfth installment of our biennial seminar series.

This iteration saw a lot of discussion on important, yet rarely analyzed topics in the domain of heuristic optimization, such as mixed-integer problems, permutation spaces, and coevolution. Moreover, it aimed at unifying existing results by discussing mathematical tools (such as drift analysis), the structure of discrete problems, and a common framework for theoretical analysis and practical implementation. Last, more recent and important topics, such as constrained and multi-objective optimization, were a major part of the seminar. We had a vivid exchange in various breakout sessions and different talks, with a great mix of junior and senior participants, which was very positively received.

Seminar June 30 – July 5, 2024 – <https://www.dagstuhl.de/24271>

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

Keywords and phrases Black-Box Optimization Heuristics, Evolution Strategies, Genetic and Evolutionary Algorithms, Runtime and Convergence Analysis, Stochastic Processes

Digital Object Identifier 10.4230/DagRep.14.6.215

1 Executive Summary

Anne Auger (Inria Saclay Ile-de-France, Palaiseau, FR)

Tobias Glasmachers (Ruhr University Bochum, DE)

Martin S. Krejca (Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, FR)

Johannes Lengler (ETH Zürich, CH)

License © Creative Commons BY 4.0 International license

© Anne Auger, Tobias Glasmachers, Martin S. Krejca, and Johannes Lengler

This seminar is part of a biennial seminar series. This year, we envisioned to focus on constraint handling, multivariate estimation-of-distribution algorithms (EDAs), as well as stochastic approximation and Markov chain stability analysis. This vision worked well for constraint handling but not so much for the other two topics, since several key invitees rejected our invitations. Nonetheless, the seminar quickly and organically refocused, and we had plenty of other important topics to discuss instead, as we detail below.

* Editor / Organizer

† Editorial Assistant / Collector



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

Theory of Randomized Optimization Heuristics, *Dagstuhl Reports*, Vol. 14, Issue 6, pp. 215–244

Editors: Anne Auger, Tobias Glasmachers, Martin S. Krejca, and Johannes Lengler



Dagstuhl Reports

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

The previous iteration of the seminar had taken place during the peak of the COVID Omicron wave, with lots of restrictions in place. We were glad to see this seminar happening again in the usual format. It was still a bit smaller than usual, due to unfortunate last-minute cancellations, but we managed to partially make up for that by inviting more young researchers. In any case, the group size of about 40 worked very well.

We used morning sessions for talks and group discussions, and afternoon sessions for breakout formats, with the exception of Monday. Several time slots were exempt from the schedule, explicitly leaving free time for individual discussions and/or leisure activities. We kept talks reasonably short, usually at a maximum of 20 minutes, leaving sufficient time for discussion. We had to stop discussions only very rarely so as to respect Dagstuhl's meal schedule. Otherwise, we managed to give sufficient space to each topic by flexibly re-scheduling a few topics on the fly.

We used the first day for group forming and to bring everyone to the same page. We started with a general introduction, a small ice-breaker game, and a series of overview talks. We had talks on optimization heuristics in discrete and continuous search spaces, which represents a classic divide in our community, as well as introductions to multi-objective optimization and constraint handling. Furthermore, we had presentations by Thomas Sauerwald on *Balls into Bins* and by Vanessa Volz on *Democratising Real-World Problem Tailored Optimisation*, two contributions a bit remote from the core topic of the seminar, so as to build bridges to neighboring fields. Kathrin Klamroth and Oswin Krause continued this series later on with presentations on *A Multiobjective Perspective on Constraint Handling* and on *Evolution in the Quantum world: on tuning quantum dot devices*.

We invited and actually nudged all young researchers to introduce themselves and their research with brief presentations. We had a total of seven such talks, taking place in the morning slots of the following days. This format was generally perceived as useful and fruitful.

Besides junior and outreach talks, there was of course also a lot of activity on core topics of the community. Two selected highlights were the presentation of Armand Gissler's proof of linear convergence of the CMA-ES algorithm on convex quadratic problems by means of Markov chain analysis, and the presentation of Per Kristian Lehre on his SLO hierarchy, a categorization of complexity classes for black-box optimization.

There was a total of 10 breakout sessions, which are all summarized in Section 4. They covered a broad spectrum of diverse and important topics, and they easily replaced the initial focus topics that we had in mind. This time, we had a fair amount of breakout sessions that garnered the attention of both the discrete and the continuous subcommunity in roughly equal parts. As such, we had sessions on negative drift (which is a tool applicable in either domain), on mixed-integer problems (which requires expertise from both domains), the general structure of discrete problems with the aim to classify them similarly to how continuous problems can be classified, and a session on an abstract framework that can be used by practitioners and analyzed by theoreticians. Moreover, we discussed rather recent hot topics such as multi-objective optimization, co-evolution, and rich surrogate models, assessing which direction to take for the future. This was complemented by a session on permutation problems, highlighting a domain that only saw little attention so far, as well as a session on how to place the theoretical analysis of randomized search heuristics within the grander spectrum of artificial intelligence. All of these sessions had a larger number of participants and vibrant discussions, showing the interest of the community in these topics.

We established a new format for collecting topics and for scheduling breakout sessions. It was inspired by a workshop that had taken place at the Lorentz Center in Leiden (Netherlands) earlier this year. Instead of collecting topics asynchronously upfront and over lunch breaks, we dedicated a short session to it. To this end, we replaced the usual classroom arrangement with a fully symmetrical setup, asking participants to step up, briefly explain their proposal, write down a title on a sheet of colored paper, and put it up on a wall. Moreover, instead of trying to reach consensus, we left it to session organizers to schedule their sessions so as to minimize (perceived) overlap. The new process was received as an improvement.

For the first time in this seminar series, we offered a trip to Trier as a social activity on Wednesday afternoon, with a hike taking place in parallel (and in the rain). This worked very smoothly, and it was a great experience, especially for participants from far away, even if they had been to Dagstuhl before.

We are very grateful for the opportunity of organizing a seminar in Dagstuhl. We have to thank for the financial support, for the comfort of rooms directly at the venue, four meals a day, wonderful facilities, and all of that including full service by very friendly, reactive, and always helpful staff. Thank you!

The organizers,

Anne Auger, Tobias Glasmachers, Martin Krejca, Johannes Lengler

2 Table of Contents

Executive Summary

Anne Auger, Tobias Glasmachers, Martin S. Krejca, and Johannes Lengler 215

Overview of Talks

Reconstructing a true individual from a noisy one: a new perspective gives new insights <i>Denis Antipov and Benjamin Doerr</i>	220
Introduction to Constrained Optimization <i>Dirk V. Arnold</i>	220
Understanding the Fitness Landscape of a Real-World Sequential Decision-Making Problem <i>Asma Atamna</i>	221
How Theoretical Considerations Can Help in Algorithm Design: 3 Examples <i>Dimo Brockhoff</i>	221
Covariance Matrix Adaptation for MCMC Sampling: some Differences with the Optimization Context <i>Alexandre Chotard</i>	221
Dumb and Good: Another Advertising Campaign For Random Parameters (Following a Power Law) <i>Benjamin Doerr</i>	222
Multiobjective optimization <i>Carlos M. Fonseca</i>	222
Surrogate methods for AL-CMA-ES with binary constraint values <i>Oskar Girardin</i>	223
Convergence proof of CMA-ES: analysis of underlying normalized Markov chains <i>Armand Gissler</i>	223
Introduction to Direct Search in Continuous Domains <i>Tobias Glasmachers</i>	224
Theoretical Aspects of Set-Quality Indicators for Multiobjective Optimization <i>Andreia Guerreiro</i>	224
Computation of a Numerical Drift and Potential Function for the (1+1)-ES <i>Alexander Jungeilges</i>	224
A Multiobjective Perspective on Constraint Handling <i>Kathrin Klamroth</i>	225
Theory of Stochastic Drift – A Guided Tour <i>Timo Kötzing</i>	225
Evolution in the Quantum world: on tuning quantum dot devices <i>Oswin Krause</i>	225
The SLO-hierarchy of pseudo-Boolean Problems <i>Per Kristian Lehre</i>	226

A Tight $O(4^k/p_c)$ Runtime Bound for a $(\mu + 1)$ GA on Jump_k for Realistic Crossover Probabilities <i>Andre Opris, Johannes Lengler, and Dirk Sudholt</i>	226
Collecting coupons: knowing when to stop <i>Jonathan Rowe</i>	227
Balls Into Bins <i>Thomas Sauerwald</i>	227
Comparison of an EDA with a hill-climber in the 1-d Ising model <i>Jonathan Shapiro</i>	227
Introduction to the Theory of Evolutionary Computation in Discrete Domains <i>Dirk Sudholt</i>	228
Crossover, Constraint Repair and Populations of Solved Subgraphs <i>Andrew M. Sutton</i>	228
Democratising Real-World Problem Tailored Optimisation <i>Vanessa Volz</i>	228
Towards Efficient and Practical Evolutionary Algorithms with Theoretical Guarantees <i>Weijie Zheng</i>	229
Working Groups	
Breakout Session: $\text{EC} \subseteq \text{Heuristic Search} \subseteq \text{AI} \implies \text{EC} \subseteq \text{AI}$? <i>Benjamin Doerr and Weijie Zheng</i>	229
Breakout Session on Theory of Multi-Objective Evolutionary Algorithms: What to Take Home From the Recent Burst of Results and Where to Go Next? <i>Benjamin Doerr and Andre Opris</i>	230
Breakout Session on Constrained Optimization <i>Carlos M. Fonseca</i>	231
Breakout Session on Mixed Integer Problems <i>Carlos M. Fonseca</i>	232
Breakout Session on Rich Surrogate Models <i>Oswin Krause</i>	233
Breakout Session on Negative Drift <i>Martin S. Krejca</i>	234
Breakout session: Coevolution <i>Per Kristian Lehre and Mario Alejandro Hevia Fajardo</i>	235
Breakout Session on the Structure of Discrete Problems <i>Jonathan Shapiro</i>	236
Breakout Session on Analysis of Permutation Problems <i>Christine Zarges</i>	237
Seminar Schedule	239
Participants	244

3 Overview of Talks

3.1 Reconstructing a true individual from a noisy one: a new perspective gives new insights

Denis Antipov (University of Adelaide, AU) and Benjamin Doerr (Ecole Polytechnique – Palaiseau, FR)

License © Creative Commons BY 4.0 International license
© Denis Antipov and Benjamin Doerr

Joint work of Denis Antipov, Benjamin Doerr, Alexandra Ivanova

Main reference Denis Antipov, Benjamin Doerr, Alexandra Ivanova: “Already Moderate Population Sizes Provably Yield Strong Robustness to Noise”, in Proc. of the Genetic and Evolutionary Computation Conference, GECCO 2024, Melbourne, VIC, Australia, July 14-18, 2024, ACM, 2024.

URL <https://doi.org/10.1145/3638529.3654196>

EAs are often used to solve problems in uncertain environments, e.g., for optimization of dynamic functions or functions with stochastic noise. They have been shown to be effective in such circumstances, however the theoretical understanding of why they are robust is very limited. In this talk, we focus on the prior noise, that is, when the individuals might be randomly changed before we evaluate their fitness. The previous works which considered this setting showed that the $(1 + 1)$ EA can withstand only low rates of noise, while population-based EAs seem to be quite robust even to a strong noise. However, the analysis of EAs on the noise is often tailored to the considered problems, which is caused by its complexity: even the elitist algorithms which are usually described by simple stochastic processes might adopt a much more complex non-elitist behavior.

In this talk, we describe a new method of analysis for prior noise problems. It is based on the observation that when we know both the parent and its offspring affected by the noise, we can estimate the distribution of the true noiseless offspring as if it was created via crossover between the parent and the noisy offspring. With this perspective, we show that the $(1 + \lambda)$ EA and the $(1, \lambda)$ EA with only a logarithmic population size are very robust even to constant-rate noise (the one which affects a constant fraction of all fitness evaluations): they can solve noisy OneMax in $O(n \log(n))$ fitness evaluations, which is the same as in the noiseless case.

This talk is a shortened version of the talk given at ThRaSH seminar earlier in June.

3.2 Introduction to Constrained Optimization

Dirk V. Arnold

License © Creative Commons BY 4.0 International license
© Dirk V. Arnold

Topics covered in the introductory talk on constrained optimization include penalty functions, Lagrange multipliers, augmented Lagrangians and the method of multipliers, and exact Lagrangian techniques. Adaptations of those methods for use in black-box optimization and different types of constraints were also discussed.

3.3 Understanding the Fitness Landscape of a Real-World Sequential Decision-Making Problem

Asma Atamna (Ruhr-Universität Bochum, DE)

License © Creative Commons BY 4.0 International license
© Asma Atamna

Real-world optimization problems can be quite different from the ones used typically to benchmark evolution strategies (ESs). They are often inherently noisy and can be highly multimodal. In this talk, we take a closer look at the fitness landscape of a sequential decision-making problem consisting of controlling a solid material-processing facility. We compare the performance of Proximal Policy Optimization (PPO) to that of the state-of-the-art ES, Covariance Matrix Adaptation Evolution Strategy (CMA-ES). Our longer-term goal is to:

- better understand what makes standard approaches, like PPO, perform better on such difficult fitness landscapes,
- draw inspiration from these standard approaches to improve the performance of ESs.

3.4 How Theoretical Considerations Can Help in Algorithm Design: 3 Examples

Dimo Brockhoff (INRIA Saclay – Palaiseau, FR)

License © Creative Commons BY 4.0 International license
© Dimo Brockhoff

Joint work of Anne Auger, Mohamed Gharafi, Nikolaus Hansen, Baptiste Plauevent Jourdain, Cheikh Touré

Progress in scientific research is typically achieved with the iterative process of observing the real world via an experiment, building a theory for the observations made, and using this theory to make predictions for a new experiment. The design of an optimization algorithms also typically follows (or should follow) this process. In my talk, I gave three examples in the context of multiobjective blackbox optimization where theoretical considerations, specific experiments, and visualizations improved the algorithm or detected failure modes.

3.5 Covariance Matrix Adaptation for MCMC Sampling: some Differences with the Optimization Context

Alexandre Chotard (Université du Littoral de la Côte d'Opale – Opal Coast, FR)

License © Creative Commons BY 4.0 International license
© Alexandre Chotard

The efficiency of a Metropolis-Hastings depends of how well suited the candidate distribution that the algorithm uses to move is with respect to the target distribution. As the algorithm explores the space of the target distribution, the information from the samples can be used to adapt the candidate distribution in the way of CMA-ES. In this talk, I will present some particularities that one must take care when doing covariance matrix adaption in a MCMC sampling context. In particular:

- Adaptation can ruin the sampling process.
- How do we measure performance?
- Differences in the stability analysis.
- The covariance learnt by CMA may not be what you want.

3.6 Dumb and Good: Another Advertising Campaign For Random Parameters (Following a Power Law)

Benjamin Doerr (*Ecole Polytechnique – Palaiseau, FR*)

License © Creative Commons BY 4.0 International license

© Benjamin Doerr

Main reference Benjamin Doerr, Huu Phuoc Le, Régis Makhlara, Ta Duy Nguyen: “Fast genetic algorithms”, in Proc. of the Genetic and Evolutionary Computation Conference, GECCO 2017, Berlin, DE, July 15-19, 2017, pp. 777–784, ACM, 2017.

URL <https://doi.org/10.1145/3071178.3071301>

This talk is a reminder that there is a simple way to choose good parameter values together with two very recent results, one mathematical and one empirical.

As result of a mathematical runtime analysis, [1] proposed the use of bit-wise mutation with a random mutation rate, chosen from a discrete power-law distribution (for bit-string representations of length n : you sample α from a discrete power-law with exponent 1.5 and flip bits independently with probability α/n). This ~~dumb~~ parameter-less mutation operator (used in the (1+1) EA) was able to optimize all jump functions in a time close to the one that is obtained with the best problem-specific mutation rate (which highly depends on the instance). Similar result were shown for other optimization problems, for more than one parameter, for other parameters, in multi-objective settings, for permutation spaces, ...

Despite these favorable results, heavy-tailed random parameter choices are still not used a lot in practice (but they are). To overcome this short-coming, and to entertain the valued audience, I shall present two recent result. (1) A mathematical runtime analysis showing yet another way how this mutation operator drastically outperforms standard bit-wise mutation (namely in filling the gaps between the optima of the subproblems used in decomposition-based multi-objective algorithms). (2) An empirical proof that heavy-tailed parameters are easy to use and can easily give better results than intensive automated parameter tuning (for the target set selection problem studied by [2]).

References

- 1 Benjamin Doerr, Huu Phuoc Le, Régis Makhlara, and Ta Duy Nguyen. Fast genetic algorithms. In *Genetic and Evolutionary Computation Conference, GECCO 2017*, pages 777–784. ACM, 2017.
- 2 Albert López Serrano and Christian Blum. A biased random key genetic algorithm applied to target set selection in viral marketing. In *Genetic and Evolutionary Computation Conference, GECCO 2022*, 241–250. ACM, 2022.

3.7 Multiobjective optimization

Carlos M. Fonseca (*University of Coimbra, PT*)

License © Creative Commons BY 4.0 International license

© Carlos M. Fonseca

Joint work of Andreia P. Guerreiro, Kathrin Klamroth, Carlos M. Fonseca

Main reference Andreia P. Guerreiro, Kathrin Klamroth, Carlos M. Fonseca: “Theoretical Aspects of Subset Selection in Multi-Objective Optimisation”, pp. 213–239, Springer, 2023.

URL https://doi.org/10.1007/978-3-031-25263-1_8


This talk provides an introduction to optimization with multiple objective functions. The concepts of Pareto dominance and solution efficiency are presented, and a number of alternative problem formulations are introduced and discussed. While the set of efficient solutions, and the corresponding non-dominated set, expose the tradeoffs between the different objectives,

selecting an overall best solution requires additional preference information to be provided by a Decision Maker, and ultimately changes the problem to be solved. Preference modelling via utility functions and outranking methods is outlined and related to scalarization approaches and population-based methods, respectively, as two important families of multiobjective optimization methods. Finally, some of the challenges posed by the experimental evaluation of such methods are identified, and some additional research topics are highlighted.

Acknowledgement. This work was partially supported by national funds through the FCT – Foundation for Science and Technology, I.P. in the scope of project CISUC—UID/CEC/00326/2020 and by the European Social Fund, through the Regional Operational Program Centro 2020.

3.8 Surrogate methods for AL-CMA-ES with binary constraint values


Oskar Girardin (Ecole Polytechnique – Palaiseau, FR)

License  Creative Commons BY 4.0 International license
© Oskar Girardin

In the context of constrained optimization problems, we might face constraints, which yield a boolean (feasible or infeasible). We tackle this issue by building a surrogate model to find an approximation of the constraint boundary. The model is used within the Augmented Lagrangian CMA-ES (AL-CMA-ES) algorithm, which is a competitive algorithm for constrained black-box optimization. In this talk we discuss the achievements as well as the challenges of our approach.

3.9 Convergence proof of CMA-ES: analysis of underlying normalized Markov chains


Armand Gissler (Ecole Polytechnique – Palaiseau, FR)

License  Creative Commons BY 4.0 International license
© Armand Gissler

In this talk, I will present the overview of a recent convergence proof of the algorithm CMA-ES (Covariance Matrix Adaptation – Evolution Strategies). More precisely, for a certain class of function called scaling-invariant, the normalization of the states of CMA-ES defines a homogeneous Markov chain, that we prove to be irreducible. When the objective function is ellipsoid (i.e., with ellipsoid level sets) we are able to establish a drift condition and the geometric ergodicity of this Markov chain. We deduce the linear convergence from a law of large numbers.

3.10 Introduction to Direct Search in Continuous Domains


Tobias Glasmachers (Ruhr-Universität Bochum, DE)

License  Creative Commons BY 4.0 International license
© Tobias Glasmachers

In this short introductory talk I introduce the main challenges and questions of interest for the analysis of evolutionary algorithms in continuous domains. One major distinction is between evolution strategies with versus without covariance matrix adaptation. While the latter class is well analyzed, the former poses significant challenges. Designing appropriate drift functions reflecting the complex dynamics of online adaptation rules for step size and covariance matrix is an open problem.

3.11 Theoretical Aspects of Set-Quality Indicators for Multiobjective Optimization

Andreia Guerreiro (INESC ID Lisboa – Lisbon, PT)

License  Creative Commons BY 4.0 International license
© Andreia Guerreiro

Set-quality indicators, which map a point set into a scalar value, are a convenient way to assess (the image of) solution sets in multiobjective optimization. Such indicators may comprise in this scalar value the proximity of the set of points to the Pareto front, as well as information regarding the distribution of points in the set. Performance assessment through quality indicators can be viewed as a transformation of the multiobjective optimization problem into a single-objective one, where the goal is to find a point set, frequently bounded in size, that maximizes the quality indicator. Consequently, each indicator is biased towards some point sets. The study of the theoretical properties of quality indicators allows to characterize the indicator-optimal subsets and, therefore, to understand such biases and their implications in performance assessment and in indicator-based evolutionary multiobjective optimization algorithms. Such theoretical aspects will be discussed in this talk.

3.12 Computation of a Numerical Drift and Potential Function for the (1+1)-ES

Alexander Jungeilges (Ruhr-Universität Bochum, DE)

License  Creative Commons BY 4.0 International license
© Alexander Jungeilges

Designing a suitable potential function is one of the crucial steps when performing drift analysis. In the context of evolution strategies, it has to capture the complex adaptation mechanism of the algorithm. Especially for evolution strategies with covariance matrix adaptation, the design of such a function imposes a significant challenge. In this flash talk, I will present a novel approach to numerically compute a suitable potential function, that provides a constant drift across a grid of parameter values.

3.13 A Multiobjective Perspective on Constraint Handling

Kathrin Klamroth (Universität Wuppertal, DE)

License © Creative Commons BY 4.0 International license
© Kathrin Klamroth

Joint work of Kathrin Klamroth, Jørgen Tind

Main reference Kathrin Klamroth, Jørgen Tind: “Constrained optimization using multiple objective programming”, *J. Glob. Optim.*, Vol. 37(3), pp. 325–355, 2007.

URL <https://doi.org/10.1007/S10898-006-9052-X>

Constraint handling techniques can often be interpreted as (partial) scalarizations of associated multiobjective optimization problems: In a first step, constraints are relaxed and re-interpreted as additional objective functions, and in a second step appropriate scalarizations are used to obtain one (or a series of) unconstrained problems. This naturally leads to the questions whether constraint handling techniques could be chosen to reflect decision maker preferences, and/or to provide additional trade-off information at little extra cost. A prominent example is the relaxation of constraints in a Lagrangian relaxation framework that corresponds to weighted sum scalarizations of associated multiobjective problems. An interesting alternative are hypervolume scalarizations that have non-linear level curves and can thus generate solution in non-convex parts of the outcome set / the Pareto front, respectively.

3.14 Theory of Stochastic Drift – A Guided Tour

Timo Kötzing (Hasso-Plattner-Institut, Universität Potsdam, DE)

License © Creative Commons BY 4.0 International license
© Timo Kötzing

URL https://hpi.de/friedrich/docs/scripts/22_AllDrift/index.html

In studying randomized search heuristics, a frequent quantity of interest is the first time a (real-valued) stochastic process obtains (or passes) a certain value. Commonly, the processes under investigation show a bias towards this goal, the stochastic drift. Turning an iteration-wise expected bias into a first time of obtaining a value is the main result of drift theorems. This thesis gives an introduction into the theory of stochastic drift, providing examples and reviewing the main drift theorems available. Furthermore, the thesis explains how these methods can be applied in a variety of contexts, including those where drift theorems seem a counter intuitive choice. Later sections examine related methods and approaches.

3.15 Evolution in the Quantum world: on tuning quantum dot devices

Oswin Krause (University of Copenhagen – Copenhagen, DK)

License © Creative Commons BY 4.0 International license
© Oswin Krause

Recently, the field of quantum computing (QC) has seen impressive results. Since quantum devices require meticulous tuning, a lot of research efforts are guided towards automating these tasks and the field is in some cases moving towards evolution. In this talk, I present some of the background of quantum devices and present some of the tuning problems that arise and how they are challenging the current state of the art in evolutionary algorithms. The two main points i discuss will be: a) as observations are often indirect, fitness evaluations

are often based on model fits, that take a lot of time and can introduce systematic errors.

b) Due to the underlying discrete state space of digital quantum computers, objective function landscapes are often a series of plateaus with the interesting physics lying on their intersections. I am happy to also host a breakout session to discuss some of these points.

3.16 The SLO-hierarchy of pseudo-Boolean Problems

Per Kristian Lehre (University of Birmingham, GB)

License © Creative Commons BY 4.0 International license

© Per Kristian Lehre

Joint work of Duc-Cuong Dang, Per Kristian Lehre

Main reference Duc-Cuong Dang, Per Kristian Lehre: “The SLO Hierarchy of pseudo-Boolean Functions and Runtime of Evolutionary Algorithms”, in Proc. of the Genetic and Evolutionary Computation Conference, GECCO 2024, Melbourne, VIC, Australia, July 14-18, 2024, ACM, 2024.

URL <https://doi.org/10.1145/3638529.3654221>

A fundamental question in theory of evolutionary computation is to understand structural properties of fitness landscapes that make them easy or hard for given classes of evolutionary algorithms. This talk introduces the so-called SLO hierarchy which provides a classification of all pseudo-Boolean functions according to the sparsity of local optima and density of fitness valleys. The hardest sub-classes of the hierarchy include problems covered by the No Free Lunch Theorem. A smaller class include problems with exponential unrestricted black-box complexity. An even smaller class cover problems with exponential elitist black box complexity. We also show a positive result, where appropriately tuned non-elitist EAs have polynomial expected runtime.

3.17 A Tight $O(4^k/p_c)$ Runtime Bound for a $(\mu + 1)$ GA on Jump_k for Realistic Crossover Probabilities

Andre Opris (Universität Passau, DE), Johannes Lengler (ETH Zürich, CH), and Dirk Sudholt (Universität Passau, DE)

License © Creative Commons BY 4.0 International license

© Andre Opris, Johannes Lengler, and Dirk Sudholt

The JUMP_k benchmark was the first problem for which crossover was proven to give a speedup over mutation-only evolutionary algorithms. Jansen and Wegener (2002) proved an upper bound of $O(\text{poly}(n) + 4^k/p_c)$ for the $(\mu+1)$ Genetic Algorithm $((\mu+1) \text{ GA})$, but only for unrealistically small crossover probabilities p_c . To this date, it remains an open problem to prove similar upper bounds for realistic p_c ; the best known runtime bound for $p_c = \Omega(1)$ is $O((n/\chi)^{k-1})$, χ a positive constant.

Using recently developed techniques, we analyse the evolution of the population diversity, measured as sum of pairwise Hamming distances, for a variant of the $(\mu+1)$ GA on Jump_k . We show that population diversity converges to an equilibrium of near-perfect diversity. This yields an improved and tight time bound of $O(\mu n \log(k) + 4^k/p_c)$ for a range of k under the mild assumptions $p_c = O(1/k)$ and $\mu \in \Omega(kn)$. For all constant k the restriction is satisfied for some $p_c = \Omega(1)$. Our work partially solves a problem that has been open for more than 20 years.

3.18 Collecting coupons: knowing when to stop

Jonathan Rowe (University of Birmingham, GB)

License © Creative Commons BY 4.0 International license
© Jonathan Rowe

Main reference Jonathan E. Rowe: “A Theoretical Investigation of Termination Criteria for Evolutionary Algorithms”, in Proc. of the Evolutionary Computation in Combinatorial Optimization – 24th European Conference, EvoCOP 2024, Held as Part of EvoStar 2024, Aberystwyth, UK, April 3-5, 2024, Proceedings, Lecture Notes in Computer Science, Vol. 14632, pp. 162–176, Springer, 2024.

URL https://doi.org/10.1007/978-3-031-57712-3_11

If you are playing the coupon collector game, but don’t know how many distinct coupons there are, how do you know when to stop playing? This problem can be seen as a model of deciding termination criteria for randomised heuristic algorithms – a subject that has received very little theoretical treatment to date.

3.19 Balls Into Bins

Thomas Sauerwald (University of Cambridge, GB)

License © Creative Commons BY 4.0 International license
© Thomas Sauerwald

Joint work of Thomas Sauerwald, Dimitrios Los

In the balanced allocation problem we wish to allocate m balls (jobs) into n bins (servers) by allowing each ball to choose from some bins sampled uniformly at random. The goal is to maintain a small gap between the maximum load and average load. For the one-choice protocol, where each ball is allocated to a random bin, the gap diverges for large m . However, for the two-choice protocol, where each ball samples two bins and is placed in the least loaded of the two, it was shown that gap is only $O(\log \log n)$ for all m . This dramatic improvement is widely known as “power of two choices”, and similar effects have been observed in hashing and routing.

In this talk, we will first give some intuition why two-choice maintains such a good balance in practice and theory. Then we will present our recent results in settings where the load information is subject to some noise. For example, the queried load information of bins might be (i) outdated, (ii) subject to some adversarial or random perturbation or (iii) only retrievable by binary queries. We prove that, roughly speaking, if the noise is not too strong, then performance is not affected and the $O(\log \log n)$ gap bound of two-choice still holds.

3.20 Comparison of an EDA with a hill-climber in the 1-d Ising model


Jonathan Shapiro (University of Manchester, GB)

License © Creative Commons BY 4.0 International license
© Jonathan Shapiro

Estimation of Distribution Algorithms replace genetic search operators with a constructed probability distribution. In this work, I compare stochastic hill-climber with a Markov Random Field model. The former has a polynomial runtime, but the later is exponential.

3.21 Introduction to the Theory of Evolutionary Computation in Discrete Domains

Dirk Sudholt (Universität Passau, DE)

License  Creative Commons BY 4.0 International license
© Dirk Sudholt

In this talk I gave a high-level overview on theory of evolutionary computation on discrete search spaces. The talk reviewed common scenarios in runtime analysis and introduced common evolutionary algorithm designs for discrete search spaces. It described recent hot topics including diversity optimisation, runtime analyses of state-of-the art multi-objective evolutionary algorithms, smart ideas for designing efficient evolutionary algorithms and analyses of complex evolutionary dynamics.

3.22 Crossover, Constraint Repair and Populations of Solved Subgraphs


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

License  Creative Commons BY 4.0 International license
© Andrew M. Sutton

In the field of combinatorial optimization, parameterized complexity theory can offer tools for understanding the influence of problem structure on hardness for evolutionary algorithms by isolating the source of problem complexity to a parameter of the input (distinct from the problem size). It can also offer insight into designing problem-class tailored operators that yield runtime guarantees in some situations. I present some recent ideas on a population-based evolutionary algorithm for solving constrained graph problems that maintains a population of feasible solutions to induced subgraphs and leverages crossover to build up feasible solutions to larger induced subgraphs. This involves two problem-specific steps during evolution: generating a special “template” parent for multi-parent uniform crossover, and an efficient constraint-repair technique inspired by the method of iterative compression from parameterized complexity theory.

3.23 Democratising Real-World Problem Tailored Optimisation


Vanessa Volz (CWI – Amsterdam, NL)

License  Creative Commons BY 4.0 International license
© Vanessa Volz

A central challenge in solving real-world optimisation problems is the lack of a systematic process. Existing approaches largely either adjust the problem representation to match a specific algorithm (e.g., by discretising) or customise an algorithm to fit the problem. However, their results tend to remain problem- or even instance-specific. The point of this talk was to start an open discussion on the challenges in abstracting real-world applications enough to conduct (theoretical) research while maintaining practical meaningfulness, based on my own experiences in empirical/industry research.

3.24 Towards Efficient and Practical Evolutionary Algorithms with Theoretical Guarantees

Wei jie Zheng (*Harbin Institute of Technology – Shenzhen, CN*)


License  Creative Commons BY 4.0 International license
© Wei jie Zheng

This talk is a self-introduction talk. It introduces my research in four aspects (algorithm choosing, parameter setting, practical problem, and practical algorithms), aiming for efficient and practical evolutionary algorithms with theoretical guarantees.

4 Working Groups

4.1 Breakout Session: $EC \subseteq \text{Heuristic Search} \subseteq AI \implies EC \subseteq AI$?

Benjamin Doerr (*Ecole Polytechnique – Palaiseau, FR*) and Wei jie Zheng (*Harbin Institute of Technology – Shenzhen, CN*)

License  Creative Commons BY 4.0 International license
© Benjamin Doerr and Wei jie Zheng

Heuristic Search is a classic subfield of artificial intelligence (AI). Clearly, evolutionary algorithms are heuristics search methods. Hence they should be part of the growing field of AI. Yet, only few evolutionary computation (EC) works appear at AI venues. This breakout session aimed at discussing this discrepancy.

Participants (in random order)

Thomas Jansen, Benjamin Doerr (session organizer), Andrew M. Sutton, Christine Zarges, Kento Uchida, Alexandre Chotard, Per Kristian Lehre, Dirk Sudholt, Jonathan L. Shapiro, Wei jie Zheng (note keeper), Denis Antipov, Sumit Adak, Asma Atamna, Stephan Frank, Oswin Krause, Tobias Glasmachers, Martin S. Krejca

Questions discussed, views and findings

- Why do we care about the relationship between evolutionary computation (EC) and artificial intelligence (AI)? Essential reason: AI is super-popular at the moment. If we are AI, we should say so, if not, then we should say what else we are.
 - We need a sufficient pool of talent for the field to develop rapidly, and AI makes it easier to establish such talent pool.
 - AI-related funding proposals are more popular and more likely to be approved, which is the basic requirement for recruiting students.
 - Industries prefer to recruit graduates who are engaged in the field of artificial intelligence, which will influence whether you can recruit enough talented students.
- What is the relationship between EC and AI? The first question should be what AI is, which is not well-defined and changes along with time. An indication that EC is generally seen as a subfield of AI is that EC papers are accepted at the big AI conferences AAAI and IJCAI and appear in the leading journals, e.g., the Artificial Intelligence Journal. EC researchers also act(ed) as associate editors of this journal, including Thomas Jansen, Christian Igel, and Benjamin Doerr.


- How narrow is the EC community in the greater AI area? Currently, the hot topic of AI is related to deep learning (deep neural network). Neural networks and evolutionary computation can be regarded as nature-inspired algorithms to some degree. Why are neural networks popular in AI, but EC is much less? Despite the similar black-box nature, neural networks have shown unbelievable improvements for specific applications.
- Concerns about AI journals and conferences, like reviewing quality.
- Views from optimization/heuristics/mathematics community. Separated communities.
- Some suggestions for EC being more visible in AI:
 - Combine the popular AI, like LLM, with EC.
 - Bayesian optimization for discrete optimization?
 - Solve practical problems that show significant improvement using EA.
 - Designable/Controllable EC considering the randomness.
 - Easy-to-use API, like CMA-ES and the ones from the deep learning community.
 - Parallel EAs.
 - More advertisement. The power of the hype behind AI can give bad optimization ideas much more light than more efficient solutions of the same problem using EC. GECCO to top AI conference?
 - More participation in AI venues as attendee, author, or reviewer.

Conclusion

Do not be shy to enter the community of AI. Become an author or reviewer in big AI conferences/journals. But maintain the characteristics and the advantages of the smaller, more coherent, and well-defined EC community.

4.2 Breakout Session on Theory of Multi-Objective Evolutionary Algorithms: What to Take Home From the Recent Burst of Results and Where to Go Next?

Benjamin Doerr (Ecole Polytechnique – Palaiseau, FR) and Andre Opris (Universität Passau, DE)

License  Creative Commons BY 4.0 International license
© Benjamin Doerr and Andre Opris

Participants. Benjamin Doerr (organizer), Timo Kötzing, Frank Neumann, Carsten Witt, Denis Antipov, Andre Opris (notekeeper), Christine Zarges, Günter Rudolph, Vanessa Volz, Sumit Adak, Weijie Zheng, Andreia P. Guerreiro, Dirk Sudholt, Martin S. Krejca

Topic. Since the first runtime analysis of the NSGA-II algorithm only two years ago, we have seen a good dozen of runtime analyses of this algorithm and variants like the NSGA-III or the SMS-EMOA. So it might be a good moment to collect what are the main insights from these results, both from the theory perspective and the possible use for practitioners, and what are the most important questions not yet understood sufficiently well.

Notes. In the session, the following observations were made by the participants.

- High relevance: Many results on top AI conferences.
- NSGA-II: First runtime analyses, approximating the Pareto front, multimodal problems, use of crossover, noise, combinatorics, critical difficulties for $m \geq 3$ objectives.
- The variants SMS-EMOA and NSGA-III often behave similarly, but not always. In particular, they perform much better for three or more objectives.


- Discussions about NSGA-III on many objectives: First theoretical insights on how to use NSGA-III on multi-objective problems, but experiments cannot be conducted: Number of reference points too high.
- SMS-EMOA: Speed-ups with stochastic population update.
- Which of these results are most useful for practitioners?
 - Negative results, especially in case of NSGA-II when dealing with three or more objectives: Limits of crowding distance.
 - Good approximation of the Pareto front when updating the crowding distance after each removal.
 - NSGA-II: Small population sizes cannot cover front.
 - How to set the reference points in an efficient way such that NSGA-III becomes more practicable? Could provide theoretical results.
- What are missing results?
 - How can we “heal” the problem with crowding distance?
Suggestion: “Density estimator” instead of crowding distance.
 - Are lower fronts of the non-dominated sorting any useful? First result on OneTrapZeroTrap where NSGA-II outperforms GSEMO when duplicates are removed.
 - Until now: NSGA-II works as good as GSEMO on deterministic benchmark functions without additional mechanisms like prohibiting duplicates. Provide benchmark function where NSGA-II beats GSEMO.
 - Understand tie breakers in the NSGA-III.
 - Overview and summary about this topic, especially NSGA-III.
 - Problems about disconnected Pareto fronts: No established benchmarks exist here. One result about functions of the form $(\sum_{i=1}^n v_i x_i, \sum_{i=1}^n w_i x_i)$. Here one could have a disconnected Pareto front since the bits are conflicting. Signs of coefficients could be positive or negative.
 - Orientate on implementations of practitioners. How to design good MOEAs?

References

- 1 Weijie Zheng, Yufei Liu, and Benjamin Doerr. A first mathematical runtime analysis of the Non-Dominated Sorting Genetic Algorithm II (NSGA-II). In *Conference on Artificial Intelligence, AAAI 2022*, pages 10408–10416. AAAI Press, 2022.

4.3 Breakout Session on Constrained Optimization

Carlos M. Fonseca (University of Coimbra, PT)

License  Creative Commons BY 4.0 International license
© Carlos M. Fonseca

Participants

Carlos, Kathrin, Dirk A., Tobias, Armand, Anne, Niko, Oskar, Dimo, Tristan, Daiki, Kento, Oswin


Summary

- Development of constraint handling in EC is partially driven by competitions.
- Continuing the discussion about Kathrin’s talk. Punchline: it is worth considering the MO perspective for additional information.

- Multi-objectivization implicitly allows to relax (otherwise hard) constraints, which can help with finding feasible points.
- It is possible (but probably not common) to aggregate all constraint violations into a single objective. This means that we get away with bi-objective optimization, which is easier than many-objective for many reasons.
- Barrier methods may be useful if staying away from the boundary initially helps to make larger steps, but not easy to do in a black-box setting.
- There does not seem to be a simple case that can be reduced to an existing theoretical analysis.
- Equality constraints versus inequality constraints: why should equality constraints be easier? Depends on the solver, in particular for active set methods.

4.4 Breakout Session on Mixed Integer Problems

Carlos M. Fonseca (University of Coimbra, PT)

License  Creative Commons BY 4.0 International license
© Carlos M. Fonseca

Participants

Carlos, Dirk A., Tobias, Anne, Niko, Asma, Oskar, Dimo, Aishwarya, Tristan, Daiki, Mario, Alexander, Kento

Summary

- Separability between discrete and continuous? Which dependency structure?
- Which test problems to consider? Is discretized ellipsoid a good problem? Maybe a moderately conditioned ellipsoid, with a max-norm neighborhood structure on the integer part. Homework for Dimo: What is the smallest condition number such that the discretized ellipsoid multi-modal?
Another rather simple yet non-separable problem: one-max plus sphere, but the optimum of the sphere is offset by the onemax value.
- Difficulties of mixed integer? With CMA-ES: Plateaus, staying unbiased despite rounding, ability to “flip only the last bit”
- Variable length, structure depending on discrete choice? Probably not in the first version.
- Test problems
 - integer
 - categorical
 - permutation
 - variable-length
 - separable?
 - adding up discrete and continuous function assumes “separability” among the discrete and continuous part.
- Inner-outer loop (inner for continuous defines fitness for outer loop optimizing discrete) is probably promising, in particular for ill-conditioned problems.
- no continuous part: what would be the main algorithm choice? only binary variables: what would be the main choice?

- performance of “older” algorithms (Emmerich / Rudolph) using discrete mutation. Li et al.: “Mixed Integer Evolution Strategies for Parameter Optimization”
- possible ways to solve mixed-integer problems (our independent “reproduction” from the paper):
 - * continuous and discrete independent with two independent loops (assuming separability)
 - * “optimal” continuous part as fitness of the discrete part (inner-outer loop idea) → probably the one way to go for some problems (where the continuous part highly depends on the discrete variables)
- question: what is the smallest condition number that the discretized ellipsoid is not multimodal (and how is this dependent on the angle)?
- what is the simplest non-separable mixed-integer function, we want to solve?

Consensus

Mixed-integer is not easy.

4.5 Breakout Session on Rich Surrogate Models

Oswin Krause (*University of Copenhagen – Copenhagen, DK*)

License © Creative Commons BY 4.0 International license
© Oswin Krause

Participants

Oswin, Dirk A., Tristan, Tobias, Vanessa, Kento, Asma, Dimo, Niko, Oskar, Daiki

Summary

- Setup: individuals in X -space, simulation $s : X \rightarrow Y$, cost function $\ell : Y \rightarrow \mathbb{R}$. We have full control over the loss ℓ by design, while the simulation s is a black box.
- Usual approach: fit surrogate (e.g., GP model) to the fitness $f = \ell \circ s$. However, this hides information, since Y may be a rich space, while the concatenation hides information by aggregating to a scalar output.
- Simplest example: s is the identity, ℓ is the sphere function. Here, if the model for s is good then we can essentially jump into the optimum.
- Another example: s is a multi-objective problem, ℓ is the hypervolume indicator. Related: ℓ is a constraint handling technique for multi-objective constraint handling.
- Complex example: s is a game level generator (e.g., a GAN), Y consists of variables related to fun, difficulty, gamer involvement and so on.
- The worst case for the a surrogate of the overall fitness is that all function values are the same, while the intermediate points differ and therefore give potentially rich information on where to move. However, in general, it is unclear whether the signal-to-noise (think of model fitting uncertainty) is better in any of the two models.

4.6 Breakout Session on Negative Drift

Martin S. Krejca (*Ecole Polytechnique – Palaiseau, FR*)

License  Creative Commons BY 4.0 International license
© Martin S. Krejca

Participants

Sumit Adak, Denis Antipov, Benjamin Doerr, Stephan Frank, Armand Gissler, Alexander Jungelges, Timo Kötzing, Oswin Krause, Martin S. Krejca, Frank Neumann, Andre Opris, Günter Rudolph, Thomas Sauerwald, Dirk Sudholt, Andrew M. Sutton, Carsten Witt, Weijie Zheng

Summary

Negative drift considers random processes $(X_t)_{t \in \mathbf{N}}$ over \mathbf{R} that, in expectation, go away from a prior determined target $a \in \mathbf{R}$ within an interval $[a, b] \subset \mathbf{R}$. The term is mostly used when speaking of negative-drift *theorems*, which aim to derive rigorous mathematical tail bounds for the process X to reach a value of at most a after a certain amount of time passed. In the theory community of randomized search heuristics, negative-drift theorems have been applied with great success for deriving exponential lower bounds in various settings, such as the maximum-matching problem [2], the NEEDLE benchmark [4], in noisy optimization [7], in fitness-proportional selection [3], or for slow fitness degradation in self-adjusting algorithms [9]. However, at least three variants of negative-drift theorems exist [4, 5, 6, 8]. This session aimed to understand to what extent these versions can be unified. A great part of the breakout group's time was spent in collecting information about these variants and finding their similarities.

All versions of the theorems require (i) negative drift and (ii) a bounded step size or exponentially declining probabilities of large jumps. And although different theorems use different tools in their proof, for example, a result by Hajek [1] or the Azuma–Hoeffding inequality, they all rely on estimating the occupation probabilities of the process. Moreover, the moment-generating function of the process seems to be crucial to the proofs. It was then noted that all current theorems require their conditions to hold for each point in time, even those that exceed the bound that one is interested in. Since the proofs mainly rely on a union bound of the occupation probabilities, this is a shortcoming that can and should be addressed. Another observation was that idle times of the process are currently not well integrated into the variants and thus require manual work.

As many randomized search heuristics are Markovian, we wondered whether such a property simplifies the assumptions of the theorems, but we did not see how. Another suggestion was to start by applying an exponentially scaled potential function to the process that results in a *positive* drift but nonetheless in an exponential expected hitting time. However, since this does not result in a tail bound directly and requires to find the correct potential function each time, this approach is more of an alternative approach than a unification.

Conclusion

We did not see how to combine the different versions easily, and we concluded that they all have their own benefits. A key take-away was that thinking more specifically about potential functions can prove useful and may forgo an application of one of the negative-drift theorems. Last, instead of having more negative-drift theorems, it might be better to have a guide on how to come up with a qualified potential function for a given process.

References

- 1 Bruce Hajek. *Hitting-time and occupation-time bounds implied by drift analysis with applications*. Advances in Applied Probability 13(3), pp. 502–525, 1982.
- 2 Oliver Giel and Ingo Wegener. *Evolutionary algorithms and the maximum matching problem*. Proc. of STACS '03, pp. 415–426, 2003.
- 3 Frank Neumann, Pietro S. Oliveto, and Carsten Witt. *Theoretical analysis of fitness-proportional selection: Landscapes and efficiency*. Proc. of GECCO '09, pp. 835–842, 2009.
- 4 Pietro S. Oliveto and Carsten Witt. *Simplified drift analysis for proving lower bounds in evolutionary computation*. Algorithmica 59(3), pp. 369–386, 2011.
- 5 Pietro S. Oliveto and Carsten Witt. *Erratum: Simplified drift analysis for proving lower bounds in evolutionary computation*. CoRR abs/1211.7184, 2012.
- 6 Timo Kötzing. *Concentration of first hitting times under additive drift*. Algorithmica 75(3), pp. 490–506, 2016.
- 7 Johannes Lengler and Ulysse Schaller. *The $(1+1)$ -EA on noisy linear functions with random positive weights*. Proc. of SSCI '18, pp. 712–719, 2018.
- 8 Johannes Lengler and Angelika Steger. *Drift analysis and evolutionary algorithms revisited*. Combinatorics, Probability & Computing 27(4), pp. 643–666, 2018.
- 9 Marc Kaufmann, Maxime Larcher, Johannes Lengler, and Xun Zou. *Self-adjusting population sizes for the $(1, \lambda)$ -EA on monotone functions*. Theoretical Computer Science 979, pp. 114181:1–114181:29, 2023.

4.7 Breakout session: Coevolution

Per Kristian Lehre (University of Birmingham, GB) and Mario Alejandro Hevia Fajardo (University of Birmingham, GB)

License © Creative Commons BY 4.0 International license

© Per Kristian Lehre and Mario Alejandro Hevia Fajardo

Main reference Per Kristian Lehre: “Runtime Analysis of Competitive Co-evolutionary Algorithms for Maximin Optimisation of a Bilinear Function”, Algorithmica, Vol. 86(7), pp. 2352–2392, 2024.

URL <https://doi.org/10.1007/S00453-024-01218-3>

Co-evolutionary algorithms have been studied empirically for a long time in Evolutionary Computation. Practitioners often report pathological behaviour, such as cycling, disengagement, and forgetting. A survey by Popovici, Bucci, Wiegand and de Jong (2012) identified runtime analysis of co-evolution as an important open problem.

The EC theory group in Birmingham has an ongoing effort in developing runtime analysis for competitive co-evolutionary algorithms. The first results include upper bounds on the runtime of the Pairwise-Dominance Co-Evolutionary Algorithm (PDCoEA) to find an ε -approximation to the maximin-solution of some instances of a bilinear problem (Lehre, GECCO2022, Algorithmica2024). Later analyses of bilinear games considered the algorithms RLS-PD (Fajardo, Lehre, Lin, GECCO2023), (μ, λ) CoEA (Fajardo & Lehre, GECCO2023), and RankDivCoEA (Fajardo & Lehre, PPSN2024). Other games have been considered, including Diagonal (a game with binary payoffs) (Lehre & Lin, PPSN 2024), as well as more complex class of symmetric zero-sum games (Benford & Lehre, GECCO 2024). Some general conclusions from this work include:


- Classical drift analysis and occupation-time bounds can be used to derive runtime. Also, a co-evolutionary level-based theorem is now available, providing upper bounds on the expected runtime of population-based CoEAs (Lehre, 2022GECCO).

- A sufficiently large population or archives can be essential. (1+1) CoEA-type algorithms forget their state (i.e., showing cycling behaviour) on several benchmark problems.
- Population diversity is often important, such as in “Games of Skill” (Benford & Lehre, GECCO24). RankDivCoEA uses a diversity mechanism which promotes unique rankings of opponents.
- Mutation rates can be critical. CoEAs have “error thresholds” (performance breaks down above critical mutation rate) similar to non-elitist EAs. However, mutation rates can be adapted automatically via self-adaptation (Fajardo et al, GECCO24).
- Traditional CoEAs do not always perform well, and it is non-trivial to design new ones. PDCoEA and RankDivCoEA have polynomial expected runtime on Bilinear, and also show excellent empirical performance on several benchmark problems.

The breakout session noted that early influential work, such as Hillis’ co-evolution of sorting networks, provided inadequate details to allow easy replication.

4.8 Breakout Session on the Structure of Discrete Problems

Jonathan Shapiro (University of Manchester, GB)

License  Creative Commons BY 4.0 International license
© Jonathan Shapiro

Participants. Andreia Guerreiro, Kathrin Klamroth, Vanessa Volz, Christine Zarges, Thomas Jansen, Per Kristian Lehre, Jon Rowe, Jon Shapiro.

Summary. Motivation was that for continuous functions there are many categories which provide global information. Examples are: $C^{(n)}$, the Lipschitz condition, and convexity. However, it seems that for functions over discrete spaces, anything can happen. Are there any such categories for discrete-space functions. They should reveal something about the function and be easy to calculate.

For example, suppose you are solving a problem which is giving continuous improvement. Could any property reveal whether it will continue to improve, or is it something like k-jump?

Earlier in the day we had seen a proposal by Per Kristian Lehre which addresses this precise issue. It is not clear that this is easy to calculate, but it does capture problems which we would consider difficult, argued Per K. Thomas J reminded us that we do not solve classes (e.g., MAXSAT). We solve instances.

What if there existed a function of an instance of a problem, which could be exponential to compute, but if you knew it, you would know that you could solve that instance in poly time. Alternatively, it could tell you where that problem is in the Lehre hierarchy.

Topics we should look at:

- Exploratory landscape analysis.
- Adversarial functions – find the most difficult instance of a class of problems.
- Constraints can hide properties of functions. We should look at constraint penalty functions.
- Black Box Complexity, but this gives worst-case result.
- Can we have a “Grey box” which reveals other information?

Proposed conclusions.

1. It is complicated.
2. We have no conclusions.

4.9 Breakout Session on Analysis of Permutation Problems

Christine Zarges (Aberystwyth University, GB)

License © Creative Commons BY 4.0 International license
© Christine Zarges

Participants (in alphabetical order)

Sumit Adak, Denis Antipov, Benjamin Doerr, Carlos M. Fonseca, Andreia P. Guerreiro, Mario Alejandro Hevia Fajardo, Thomas Jansen, Kathrin Klamroth, Timo Kötzing, Martin S. Krejca, Per Kristian Lehre, Frank Neumann, Andre Opris, Aishwarya Radhakrishnan, Jonathan Rowe, Günter Rudolph, Jonathan L. Shapiro, Dirk Sudholt, Andrew M. Sutton, Carsten Witt, Christine Zarges, Weijie Zheng

Summary

For many real-world problems, permutations constitute a natural way to represent solutions. Typical examples include problems where the order of elements (e.g., scheduling) or adjacency (e.g., travelling salesperson) matter [4]. A series of special sessions and workshops investigated different aspects specific to permutation problems¹. However, in particular with respect to runtime analysis and fixed budget analysis, permutation problems have so far received much less attention in comparison to pseudo-Boolean or continuous problems.

As a starting point, this breakout session discussed a small (incomplete) selection of existing work: Rowe et al. [8] consider which crossover operators are useful for which representations while Irurozki [5] investigates various mutation operators and the distance measures they imply. Schiavinotto and Stützle [10] present a review of metrics on permutations for search landscape analysis. One of the first runtime analyses of permutation problems considered sorting with different measures of sortedness [9]. More recently, Doerr et al. [3] propose a general way to construct permutation benchmarks from classic pseudo-Boolean benchmarks, and Bassin and Buzdalov [1] discuss how to adapt the $(1 + (\lambda, \lambda))$ GA to permutation spaces. A few real-world permutation problems, e.g., graph drawing [2] and variable ordering in ordered binary decision diagrams [6], have also been analysed.

Throughout the discussion various problems were mentioned that could be considered in future work, e.g., Bin Packing, Graph Colouring, Knapsack, (weighted) variants of the Travelling Salesperson Problem, and linear/quadratic assignment problems. For some of these problems, a permutation can be used as input for a greedy algorithm. During the discussion, questions were raised on how more general benchmarks could be obtained and how relevant problems classes could be defined as permutation problems are often NP-hard.

Dealing with permutations directly instead of using “decoder functions” implies that mutation and crossover operators need to be designed in a way that preserves the permutation property of a search point. Parts of the discussion considered potential analyses of existing decoder functions, particularly those using real-valued representations, to facilitate the analysis of crossover. Other points of discussion included geometric crossover for permutations [7] and partition crossover [11].

¹ http://www.sc.ehu.es/ccwbayes/cec2022_ecperm/index.html

Summary

There seemed to be a general consensus that we currently only have little understanding on working principles of mutation and crossover operators for permutation problems and that more research in this direction is needed. However, concrete next steps to achieve this were less obvious in the discussion. It would be nice to identify “classes” of permutation problems and establish a framework that maps operators to these problem classes. One way to achieve this could be to concentrate on operators and the fitness landscapes they imply. Moreover, none of the participants was aware of existing runtime analyses considering crossover operators for permutations.

References

- 1 A. O. Bassin and M. Buzdalov. The $(1 + (\lambda, \lambda))$ genetic algorithm for permutations. In C. A. C. Coello, editor, *GECCO '20: Genetic and Evolutionary Computation Conference, Companion Volume, Cancún, Mexico, July 8-12, 2020*, pages 1669–1677. ACM, 2020.
- 2 J. Baumann, I. Rutter, and D. Sudholt. Evolutionary computation meets graph drawing: Runtime analysis for crossing minimisation on layered graph drawings. In X. Li and J. Handl, editors, *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO 2024, Melbourne, VIC, Australia, July 14-18, 2024*. ACM, 2024.
- 3 B. Doerr, Y. Ghannane, and M. I. Brahim. Runtime analysis for permutation-based evolutionary algorithms. *Algorithmica*, 86(1):90–129, 2024.
- 4 A. E. Eiben and J. E. Smith. *Introduction to Evolutionary Computing, Second Edition*. Natural Computing Series. Springer, 2015.
- 5 E. Irurozki. *Sampling and learning distance-based probability models for permutation spaces*. Phd thesis, University of the Basque Country, 2014. Available at https://github.com/isg-ehu/PhD-Dissertations/blob/master/2014_phd_ekhine_irurozki.pdf.
- 6 T. Jansen and C. Zarges. Experimental and theoretical analysis of local search optimising OBDD variable orderings. In T. Stützle and M. Wagner, editors, *Evolutionary Computation in Combinatorial Optimization – 24th European Conference, EvoCOP 2024, Held as Part of EvoStar 2024, Aberystwyth, UK, April 3-5, 2024, Proceedings*, volume 14632 of *Lecture Notes in Computer Science*, pages 177–192. Springer, 2024.
- 7 A. Moraglio and R. Poli. Geometric crossover for the permutation representation. *Intelligenza Artificiale*, 5(1):49–63, 2011.
- 8 J. E. Rowe, M. D. Vose, and A. H. Wright. Group properties of crossover and mutation. *Evol. Comput.*, 10(2):151–184, 2002.
- 9 J. Scharnow, K. Tinnefeld, and I. Wegener. The analysis of evolutionary algorithms on sorting and shortest paths problems. *J. Math. Model. Algorithms*, 3(4):349–366, 2004.
- 10 T. Schiavinotto and T. Stützle. A review of metrics on permutations for search landscape analysis. *Comput. Oper. Res.*, 34(10):3143–3153, 2007.
- 11 R. Tinós, L. D. Whitley, and G. Ochoa. A new generalized partition crossover for the traveling salesman problem: Tunneling between local optima. *Evol. Comput.*, 28(2):255–288, 2020.

5 Seminar Schedule

Monday, July 01, 2024 – Lecture Hall Saarbrücken

– 09:00	Breakfast
09:00 – 09:15	Welcome and seminar opening
09:15 – 10:00	Ice-breaking
10:00 – 10:30	Coffee break
10:30 – 11:15	Introductory talk: Daiki Morinaga and Tobias Glasmachers on <i>Continuous Domains</i>
11:15 – 12:00	Introductory talk: Thomas Sauerwald on <i>Balls Into Bins</i>
12:15 – 13:30	Lunch
13:30 – 14:15	Time for individual discussions
14:15 – 15:00	Introductory talk: Carlos Fonseca on <i>Multiobjective Optimization</i>
15:00 – 15:30	Introductory Talk: Dirk Sudholt on <i>Discrete Domains</i>
15:30 – 16:00	Cake
16:00 – 16:45	Introductory talk: Dirk Arnold on <i>Constraint Handling</i>
16:45 – 16:55	Vanessa Volz
16:55 – 17:00	Timo Kötzing on his habilitation thesis
18:00 – 19:00	Dinner
19:00 –	Individual discussions, soccer, cheese platter, drinks, and games

Tuesday, July 02, 2024 – Lecture Hall Saarbrücken

– 09:00	Breakfast
9:15 – 10:05	Junior flash talks: speakers on <i>topics</i>
9:15 – 9:25	Tristan Marty introducing himself
9:25 – 9:35	Aishwarya Radhakrishnan introducing herself
9:35 – 9:45	Alexander Jungeilges introducing himself
9:45 – 9:55	Mario Alejandro Hevia Fajardo introducing himself
10:10 – 10:35	Coffee break
10:35 – 10:45	Stephan Frank introducing himself
10:55 – 11:15	Armand Gissler on <i>Convergence proof of CMA-ES: analysis of underlying normalized Markov chains</i>
11:25 – 11:45	Per Kristian Lehre on <i>The SLO Hierarchy</i>
12:15 – 13:30	Lunch
14:00 – 14:15	Organization of breakout sessions
14:15 – 15:30	Breakout Sessions: <ul style="list-style-type: none"> ◊ Negative Drift ◊ Characterization of Discrete Problems ◊ Mixed Integer
15:30 – 16:00	Cake
16:00 – 16:20	Kathrin Klamroth on <i>A Multiobjective Perspective on Constraint Handling</i>
16:30 – 17:45	Breakout Sessions: <ul style="list-style-type: none"> ◊ Constrained Optimization ◊ Theory of Multi-Objective Evolutionary Algorithms ◊ Coevolution
18:00 – 19:00	Dinner
20:00 –	Individual discussions, soccer, cheese platter, drinks

Wednesday, July 03, 2024 – Lecture Hall Saarbrücken

– 09:00	Breakfast
9:00 – 9:10	Oskar Girardin introducing himself
9:20 – 9:35	Weijie Zheng introducing himself
9:45 – 10:00	Oswin Krause on <i>Evolution in the Quantum world: on tuning quantum dot devices</i>
10:10 – 10:50	Coffee break
10:50 – 11:10	Dimo Brockhoff on <i>How Theoretical Considerations Can Help in Algorithm Design: 3 Examples</i>
11:20 – 11:40	Andre Opris on <i>A Tight $O(4^k/p_c)$ Runtime Bound for a $(\mu+1)$ GA on Jump_k for Realistic Crossover Probabilities</i>
12:15 – 13:30	Lunch
13:30 – 15:30	Social event H: beautiful hike in acceptable weather
14:00 – 18:30	Social event T: trip to Trier
15:30 – 16:00	Cake
16:00 – 18:30	Free time
18:30 – 19:30	Dinner
19:30 –	Individual discussions, preparing my talk tomorrow, beer, no soccer

Thursday, July 04, 2024 – Lecture Hall Saarbrücken

– 09:00	Breakfast
9:00 – 9:20	Benjamin Doerr on <i>Dumb and Good: Another Advertising Campaign For Random Parameters (Following a Power Law)</i>
9:30 – 9:50	Alexandre Chotard on <i>Covariance Matrix Adaptation for MCMC Sampling: some Differences with the Optimization Context</i>
10:00 – 10:35	Coffee break
10:35 – 10:55	Denis Antipov on <i>Reconstructing a true individual from a noisy one: a new perspective gives new insights</i>
11:05 – 11:25	Jonathan L. Shapiro on <i>Comparison of an EDA with a hill-climber in the 1-d Ising model</i>
11:35 – 11:45	Carlos M. Fonseca : ROAR-NET COST Action
12:15 – 13:30	Lunch
13:55 – 14:10	Group photo ; we meet in front of the castle
14:15 – 15:30	Breakout Sessions: <ul style="list-style-type: none"> ◊ Permutation Search Spaces ◊ Rich Surrogate Models
15:30 – 16:00	Cake
16:30 – 17:45	Breakout Sessions: <ul style="list-style-type: none"> ◊ Abstractions for theoretical analysis and practical implementation ◊ $EC \subseteq HS \subseteq AI$
18:00 – 19:00	Dinner
20:00 –	Party : traditional wine and cheese platter

Friday, July 05, 2024 – Lecture Hall Saarbrücken

– 09:00	Breakfast
9:00 – 9:15	Asma Atamna on <i>Understanding the Fitness Landscape of a Real-World Sequential Decision-Making Problem</i>
9:25 – 9:45	Andreia P. Guerreiro on <i>Theoretical Aspects of Set-Quality Indicators for Multiobjective Optimization</i>
9:55 – 10:30	Coffee break
10:30 – 10:50	Jonathan Rowe on <i>Collecting coupons: knowing when to stop</i>
11:00 – 11:20	Andrew M. Sutton on <i>Crossover, Constraint Repair and Populations of Solved Subgraphs</i>
11:30 – 12:00	Closing session, feedback, ideas
12:15 – 13:30	Lunch
13:30 –	Departure

Participants

- Sumit Adak
Technical University of Denmark
– Lyngby, DK
- Denis Antipov
University of Adelaide, AU
- Dirk Arnold
Dalhousie University –
Halifax, CA
- Asma Atamna
Ruhr-Universität Bochum, DE
- Anne Auger
INRIA Saclay – Palaiseau, FR
- Dimo Brockhoff
INRIA Saclay – Palaiseau, FR
- Alexandre Chotard
Calais University, FR
- Benjamin Doerr
Ecole Polytechnique –
Palaiseau, FR
- Carlos M. Fonseca
University of Coimbra, PT
- Stephan Frank
Ruhr-Universität Bochum, DE
- Oskar Girardin
Ecole Polytechnique –
Palaiseau, FR
- Armand Gissler
Ecole Polytechnique –
Palaiseau, FR
- Tobias Glasmachers
Ruhr-Universität Bochum, DE
- Andreia P. Guerreiro
INESC-ID – Lisbon, PT
- Nikolaus Hansen
Inria & Institut Polytechnique de
Paris – Palaiseau, FR
- Mario Alejandro Hevia Fajardo
University of Birmingham, GB
- Thomas Jansen
Aberystwyth University, GB
- Alexander Jungeilges
Ruhr-Universität Bochum, DE
- Kathrin Klamroth
Universität Wuppertal, DE
- Timo Kötzing
Hasso-Plattner-Institut,
Universität Potsdam, DE
- Oswin Krause
University of Copenhagen, DK
- Martin S. Krejca
Ecole Polytechnique –
Palaiseau, FR
- Per Kristian Lehre
University of Birmingham, GB
- Tristan Marty
Ecole Polytechnique –
Palaiseau, FR
- Daiki Morinaga
University of Tsukuba, JP
- Frank Neumann
University of Adelaide, AU
- Andre Opris
Universität Passau, DE
- Aishwarya Radhakrishnan
Hasso-Plattner-Institut,
Universität Potsdam, DE
- Jonathan Rowe
University of Birmingham, GB
- Günter Rudolph
TU Dortmund, DE
- Thomas Sauerwald
University of Cambridge, GB
- Jonathan L. Shapiro
University of Manchester, GB
- Dirk Sudholt
Universität Passau, DE
- Andrew M. Sutton
University of Minnesota –
Duluth, US
- Kento Uchida
Yokohama National
University, JP
- Vanessa Volz
CWI – Amsterdam, NL
- Carsten Witt
Technical University of Denmark
– Lyngby, DK
- Christine Zarges
Aberystwyth University, GB
- Weijie Zheng
Harbin Institute of Technology –
Shenzhen, CN

