



DAGSTUHL REPORTS

Volume 12, Issue 2, February 2022

Logic and Random Discrete Structures (Dagstuhl Seminar 22061) <i>Erich Grädel, Phokion G. Kolaitis, Marc Noy, and Matthias Naaf</i>	1
Computation and Reconfiguration in Low-Dimensional Topological Spaces (Dagstuhl Seminar 22062) <i>Maïke Buchin, Anna Lubiw, Arnaud de Mesmay, Saul Schleimer, and Florestan Brunck</i>	17
New Perspectives in Symbolic Computation and Satisfiability Checking (Dagstuhl Seminar 22072) <i>Erika Abraham, James H. Davenport, Matthew England, and Alberto Griggio</i>	67
Theory of Randomized Optimization Heuristics (Dagstuhl Seminar 22081) <i>Anne Auger, Carlos M Fonseca, Tobias Friedrich, Johannes Lengler, and Armand Gïssler</i>	87
Deep Learning and Knowledge Integration for Music Audio Analysis (Dagstuhl Seminar 22082) <i>Rachel Bittner, Meinard Müller, Juhan Nam, Michael Krause, and Yigitcan Özer</i>	103
AI for the Social Good (Dagstuhl Seminar 22091) <i>Claudia Clopath, Ruben De Winne, and Tom Schaul</i>	134

ISSN 2192-5283

Published online and open access by

Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at <https://www.dagstuhl.de/dagpub/2192-5283>

Publication date

August, 2022

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <https://dnb.d-nb.de>.

License

This work is licensed under a Creative Commons Attribution 4.0 International license (CC BY 4.0).



In brief, this license authorizes each and everybody to share (to copy, distribute and transmit) the work under the following conditions, without impairing or restricting the authors' moral rights:

- Attribution: The work must be attributed to its authors.

The copyright is retained by the corresponding authors.

Aims and Scope

The periodical *Dagstuhl Reports* documents the program and the results of Dagstuhl Seminars and Dagstuhl Perspectives Workshops.

In principal, for each Dagstuhl Seminar or Dagstuhl Perspectives Workshop a report is published that contains the following:

- an executive summary of the seminar program and the fundamental results,
- an overview of the talks given during the seminar (summarized as talk abstracts), and
- summaries from working groups (if applicable).

This basic framework can be extended by suitable contributions that are related to the program of the seminar, e. g. summaries from panel discussions or open problem sessions.

Editorial Board

- Elisabeth André
- Franz Baader
- Daniel Cremers
- Goetz Graefe
- Reiner Hähnle
- Barbara Hammer
- Lynda Hardman
- Oliver Kohlbacher
- Steve Kremer
- Rupak Majumdar
- Heiko Mantel
- Albrecht Schmidt
- Wolfgang Schröder-Preikschat
- Raimund Seidel (*Editor-in-Chief*)
- Heike Wehrheim
- Verena Wolf
- Martina Zitterbart

Editorial Office

Michael Wagner (*Managing Editor*)
Michael Didas (*Managing Editor*)
Jutka Gasiorowski (*Editorial Assistance*)
Dagmar Glaser (*Editorial Assistance*)
Thomas Schillo (*Technical Assistance*)

Contact

Schloss Dagstuhl – Leibniz-Zentrum für Informatik
Dagstuhl Reports, Editorial Office
Oktavie-Allee, 66687 Wadern, Germany
reports@dagstuhl.de
<https://www.dagstuhl.de/dagrep>

Digital Object Identifier: 10.4230/DagRep.12.2.i

Logic and Random Discrete Structures

Erich Grädel^{*1}, Phokion G. Kolaitis^{*2}, Marc Noy^{*3}, and Matthias Naaf^{†4}

1 RWTH Aachen, DE. graedel@logic.rwth-aachen.de

2 University of California Santa Cruz and IBM Research, US. kolaitis@ucsc.edu

3 UPC Barcelona Tech, ES. marc.noy@upc.edu

4 RWTH Aachen, DE. naaf@logic.rwth-aachen.de

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 22061 “Logic and Random Discrete Structures”. The main topic of this seminar has been the analysis of large random discrete structures, such as trees, graphs, or permutations, from the perspective of mathematical logic. It has brought together both experts and junior researchers from a number of different areas where logic and random structures play a role, with the goal to establish new connections between such areas and to encourage interactions between foundational research and different application areas, including probabilistic databases.

Seminar February 6–11, 2022 – <http://www.dagstuhl.de/22061>

2012 ACM Subject Classification Theory of computation → Logic; Theory of computation → Computational complexity and cryptography; Mathematics of computing → Discrete mathematics

Keywords and phrases combinatorics, complexity theory, logic, random structures, probabilistic databases

Digital Object Identifier 10.4230/DagRep.12.2.1

1 Executive Summary

Erich Grädel

Phokion G. Kolaitis

Marc Noy

License  Creative Commons BY 4.0 International license
© Erich Grädel, Phokion G. Kolaitis, and Marc Noy

Topic and Goals of the Seminar

The analysis of large random discrete structures, such as trees, graphs, or permutations, is a focus of research in contemporary discrete mathematics. Logic provides a useful and powerful formalism for expressing and classifying discrete structures; moreover, it is intimately linked to the study of algorithms, computational complexity, and structural graph theory. Over the past several decades, researchers have studied random discrete structures from a logical perspective. The first significant result in this direction was the zero-one law for first-order logic under the uniform measure; this seminal result, was followed by the discovery of ‘logical limit laws’ or ‘convergence laws’ for several different models of random discrete structures and for various logics of significance in computer science. In more recent years, a renewed impetus has emerged for research activity on random discrete structures from a logical

* Editor / Organizer

† Editorial Assistant / Collector



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

Logic and Random Discrete Structures, *Dagstuhl Reports*, Vol. 12, Issue 2, pp. 1–16

Editors: Erich Grädel, Phokion G. Kolaitis, Marc Noy, and Matthias Naaf



Dagstuhl Reports

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

perspective. This is in part due to the availability of new methods and techniques, including asymptotic enumeration, discrete harmonic analysis, an extension of Gowers norms, and limit structures. Exciting new results on random geometric graphs, graphs on surfaces, classes of sparse graphs, graph limits, and flag algebras have been established. On the computer science side, there has been a systematic exploration of probabilistic databases, which has brought together databases, logic, and random structures. The main aim of this seminar has been to bring together some of the foremost experts from these different fields, as well as junior researchers who may become motivated to work deeper in the frontier of logic and random structures. In addition to making tangible progress on some of the currently outstanding open problems in this area, we wanted to establish new connections between (classical) random discrete structures, flag algebras, and sparse graph limits, both in terms of identifying new research questions and embarking on new collaborations, as well as fruitful interaction between foundational research and different application areas, including probabilistic databases.

Organisation and Activities

Despite the restrictions and problems caused by corona pandemic, the seminar had originally been intended as a non-hybrid event with all participants on site. At the end, however, this turned out to be infeasible; as a result, two of the invited survey talks and a number of the contributed talks had to be given remotely via Zoom.

The organisers created a schedule consisting of four invited one-hour survey talks, and more focussed regular contributions proposed by the participants. The survey talks were given by

- Albert Atserias on certifying the chromatic number of a random graph;
- Fiona Skerman on the inference of underlying community structures in partially observed graphs;
- Dan Suciú on probabilistic databases;
- Patrice Ossona de Mendez on limits of graphs.

The talks of Fiona Skerman and Patrice Ossona de Mendez were given over Zoom.

In addition, there were 18 contributed talks, 11 of which were given on site, and 7 remotely via Zoom.

Overall, the organisers regard the seminar to have been a very successful scientific event. There was a general view shared by all participants that the community working on logic and random structures is in excellent shape, with interesting new developments and exciting results in many different directions. The participants clearly expressed the wish to have a future meeting of this community, be in in Dagstuhl or elsewhere, within the next two to three years.

The organisers are grateful to the Scientific Directorate and to the staff of the Center for their support of this seminar.

2 Table of Contents

Executive Summary

<i>Erich Grädel, Phokion G. Kolaitis, and Marc Noy</i>	1
--	---

Overview of Talks

On certifying the chromatic number of a random graph <i>Albert Atserias</i>	5
The almost-sure theories of classes defined by forbidden homomorphisms <i>Manuel Bodirsky</i>	5
The modal logic of almost sure frame validities in the finite <i>Valentin Goranko</i>	6
Limit laws for existential monadic second order logic <i>Maksim Zhukovskii</i>	6
Improved bounds for acyclic coloring parameters <i>Lefteris M. Kirousis</i>	7
Quasirandom combinatorial structures <i>Daniel Král'</i>	7
Partially observing graphs – when can we infer underlying community structure? <i>Fiona Skerman</i>	8
Heavy-tailed distributions for random SAT <i>Andrei A. Bulatov</i>	8
Convergence law for random permutations <i>Valentin Féray</i>	9
Logic and property testing on graphs of bounded degree <i>Isolde Adler</i>	9
Limiting probabilities of first order properties of random sparse graphs <i>Marc Noy</i>	10
Logical limit laws for Mallows random permutations <i>Tobias Müller</i>	10
On first order model checking on Erdős-Rényi graphs <i>Peter Rossmanith</i>	10
Probabilistic databases – overview <i>Dan Suciu</i>	11
Explaining query answers through probabilistic databases <i>Benny Kimelfeld</i>	11
Probabilistic query evaluation with bag semantics <i>Peter Lindner</i>	12
Limits of graphs <i>Patrice Ossona de Mendez</i>	12
Zero-one laws and almost sure valuations of first-order logic in semiring semantics <i>Matthias Naaf</i>	13

Moser-Tardos algorithm with small number of random bits <i>Oleg Pikhurko</i>	13
The repeated insertion model (RIM): probabilistic inference and applications <i>Batya Kenig</i>	14
Repairs, measures, and complexity for constraints violations in databases <i>Sudeepa Roy</i>	14
Ordered graphs of bounded twin-width <i>Szymon Torunczyk</i>	15
Participants	16
Remote Participants	16

3 Overview of Talks

The talks are listed in the order of presentation during the seminar.

3.1 On certifying the chromatic number of a random graph

Albert Atserias (UPC Barcelona Tech, ES)

License  Creative Commons BY 4.0 International license
© Albert Atserias

A standard first-moment calculation shows that, for every fixed integer $q > 1$, the chromatic number $\chi(G)$ of an Erdős-Rényi random graph G of any sufficiently large constant density $d > c(q)$ of edges is, asymptotically almost surely, larger than q . We study the question whether there exist efficiently checkable certificates for the statement “ $\chi(G) > q$ ” that apply to such a random graph with high probability. First we overview the known negative results that show that bounded local consistency methods do not suffice. Then we ask whether semi-definite programming methods suffice and provide some new observations indicating that they do. This is in sharp contrast to the case of random k -CNF formulas where it is known that not even semi-definite programming methods suffice to certify their almost sure unsatisfiability at any constant density of clauses to variables.

3.2 The almost-sure theories of classes defined by forbidden homomorphisms

Manuel Bodirsky (TU Dresden, DE)

License  Creative Commons BY 4.0 International license
© Manuel Bodirsky

Joint work of Manuel Bodirsky, Colin Jahel

Main reference Manuel Bodirsky, Colin Jahel: “Asymptotic Theories of Classes Defined by Forbidden Homomorphisms”, CoRR, Vol. abs.2204.01404, 2022.

URL <https://arxiv.org/abs/2204.01404>

This talk is about the almost-sure theories for classes of finite structures that are specified by homomorphically forbidding a finite set \mathcal{F} of finite structures. If \mathcal{F} consists of undirected graphs, a full description of these theories can be derived from the Kolaitis-Proemel-Rothschild theorem, which treats the special case where $\mathcal{F} = K_n$. The corresponding question for finite sets \mathcal{F} of directed graphs is wide open. We present a description of the almost-sure theories of classes described by homomorphically forbidding finite sets \mathcal{F} of oriented trees.

3.3 The modal logic of almost sure frame validities in the finite

Valentin Goranko (University of Stockholm, SE)

License © Creative Commons BY 4.0 International license
© Valentin Goranko

Main reference Valentin Goranko: “The Modal Logic of Almost Sure Frame Validities in the Finite”, in Proc. of the 13th Conference on Advances in Modal Logic, AiML 2020, Helsinki, Finland, August 24-28, 2020, pp. 249–268, College Publications, 2020.

URL <http://www.aiml.net/volumes/volume13/Goranko.pdf>

A modal formula is almost surely frame-valid in the finite if the probability that it is valid in a randomly chosen finite frame with n states is asymptotically 1 as n grows unboundedly. In this talk I discuss the normal modal logic ML^{as} of all modal formulae that are almost surely frame-valid in the finite. Because of the failure of the zero-one law for frame validity in modal logic, the logic ML^{as} extends properly the modal logic of the countable random frame ML^f , which was completely axiomatized in a 2003 paper by Goranko and Kapron. Thus, the logic ML^{as} is still relatively little known and understood.

In this talk I first introduce the logic ML^{as} , present what is known about it, including the so far known axioms coming from ML^f and a certain model-theoretic characterisation of its additional validities beyond those in ML^f . I then raise some open problems and conjectures regarding the missing additional axioms over ML^f and the explicit description of the complete axiomatisation of ML^{as} , which may turn out to hinge on difficult combinatorial-probabilistic arguments and calculations.

3.4 Limit laws for existential monadic second order logic

Maksim Zhukovskii (MIPT – Dolgoprudny, RU)

License © Creative Commons BY 4.0 International license
© Maksim Zhukovskii

The classical result of Glebskii, Kogan, Liogon’kii, Talanov (1969) and of Fagin (1976) states that the dense (the probability of appearance of an edge p is constant) binomial random graph $G(n, p)$ obeys first order 0-1 law. On the other hand, Kaufmann and Shelah (1985) proved that there exists a monadic second order (MSO) sentence such that its truth probability on $G(n, p)$ does not converge. In 2001, Le Bars proved the same non-convergence result for the existential fragment of MSO. In the talk, I am going to review limit laws for existential monadic second order (EMSO) properties of binomial random graphs both in dense and sparse regimes. In particular, I will state our recent result that disproves the conjecture of Le Bars: there exists an EMSO sentence with 2 first order variables such that its truth probability does not converge.

3.5 Improved bounds for acyclic coloring parameters

Lefteris M. Kirousis (University of Athens, GR)

License © Creative Commons BY 4.0 International license
© Lefteris M. Kirousis

Joint work of Lefteris M. Kirousis, John Livieratos

The *acyclic chromatic number* of a graph is the least number of colors needed to properly color its vertices so that none of its cycles has only two colors. The *acyclic chromatic index* is the analogous graph parameter for edge colorings. We first show that the acyclic chromatic index is at most $2\Delta - 1$, where Δ is the maximum degree of the graph. We then show that for all $\epsilon > 0$ and for Δ large enough (depending on ϵ), the acyclic chromatic number of the graph is at most $\lceil (2^{-1/3} + \epsilon)\Delta^{4/3} \rceil + \Delta + 1$. Both results improve long chains of previous successive advances. Both are algorithmic, in the sense that the colorings are generated by randomized algorithms. However, in contrast with extant approaches, where the randomized algorithms assume the availability of enough colors to guarantee properness deterministically, and use additional colors for randomization in dealing with the bichromatic cycles, our algorithms may initially generate colorings that are not necessarily proper; they only aim at avoiding cycles where all pairs of edges, or vertices, that are one edge, or vertex, apart in a traversal of the cycle are homochromatic (of the same color). When this goal is reached, they check for properness and if necessary they repeat until properness is attained.

3.6 Quasirandom combinatorial structures

Daniel Král' (Masaryk University – Brno, CZ)

License © Creative Commons BY 4.0 International license
© Daniel Král'

Joint work of Timothy Chan, Jacob Cooper, Robert Hancock, Adam Kabelá, Daniel Král', Ander Lamaison, Taísa Martins, Samuel Mohr, Jonathan Noel, Roberto Parente, Yanitsa Pehova, Oleg Pikhurko, Maryam Sharifzadeh, Fiona Skerman, Jan Volec

A combinatorial structure is said to be quasirandom if it resembles a random structure in a certain robust sense. The notion of quasirandom graphs, developed in the work of Rödl, Thomason, Chung, Graham and Wilson in 1980s, is particularly robust as several different properties of truly random graphs, e.g., subgraph density, uniform edge distribution and spectral properties, are satisfied by a large graph if and only if one of them is.

We will discuss how the classical results on quasirandom graphs can be viewed through the lenses of the theory of combinatorial limits and apply the tools offered by this theory to study quasirandomness of tournaments, permutations and Latin squares. We show that the same phenomenon as in the case of graphs, when quasirandomness is captured by the densities of finitely many substructures (in the case of graphs, an edge and any even cycle), appears in relation to these structures, too. We also discuss what minimal sets of substructures have this property and give characterizations of such sets in several of the considered settings.

References

- 1 Timothy Chan, Daniel Král', Jonathan Noel, Yanitsa Pehova, Maryam Sharifzadeh, and Jan Volec. Characterization of quasirandom permutations by a pattern sum. *Random Structures & Algorithms*, 57(4):920–939, 2020. doi:10.1002/rsa.20956.
- 2 Jacob Cooper, Daniel Král', Ander Lamaison, and Samuel Mohr. Quasirandom latin squares. *Random Structures & Algorithms*, 2020, doi:https://doi.org/10.1002/rsa.21060.

- 3 Robert Hancock, Adam Kabela, Daniel Král', Taísa Martins, Roberto Parente, Fiona Skerman, and Jan Volec. No additional tournaments are quasirandom-forcing. [arXiv:1912.04243](https://arxiv.org/abs/1912.04243), 2019.
- 4 Daniel Král' and Oleg Pikhurko. Quasirandom permutations are characterized by 4-point densities. *Geometric and Functional Analysis*, 23(2):570–579, 2013. doi:10.1007/s00039-013-0216-9.

3.7 Partially observing graphs – when can we infer underlying community structure?

Fiona Skerman (Uppsala University, SE)

License © Creative Commons BY 4.0 International license
© Fiona Skerman

Joint work of Colin McDiarmid, Fiona Skerman.

Main reference Colin McDiarmid, Fiona Skerman: “Modularity and edge sampling”, CoRR, Vol. abs/2112.13190, 2021.

URL <https://arxiv.org/abs/2112.13190>

Suppose edges in an underlying graph G appear independently with some probability in our observed graph G' – or alternately that we can query uniformly random edges. We describe how high a sampling probability we need to infer the modularity of the underlying graph.

Modularity is a function on graphs which is used in algorithms for community detection. For a given graph G , each partition of the vertices has a modularity score, with higher values indicating that the partition better captures community structure in G . The (max) modularity $q^*(G)$ of the graph G is defined to be the maximum over all vertex partitions of the modularity score, and satisfies $0 \leq q^*(G) \leq 1$.

In the seminar I will spend time on intuition for the behaviour of modularity, how it can be approximated, links to other graph parameters and to present some open problems.

3.8 Heavy-tailed distributions for random SAT

Andrei A. Bulatov (Simon Fraser University – Burnaby, CA)

License © Creative Commons BY 4.0 International license
© Andrei A. Bulatov

Joint work of Andrei A. Bulatov, Oleksii Omelchenko

While the focus of research on Random SAT have been on uniformly distributed instances, several models were suggested, in which the distribution from which instances are sampled is far from uniform. Sometimes such models are motivated by analyzing ‘practical’ distributions, sometimes they appear to be mathematically natural and sound. We study one such model, the Configuration Model, which is parametrized by a distribution of degrees of variables. To sample from this model we start with a set of variables, then for each of them we sample a degree and create the prescribed number of copies. Each copy (or clone) is then negated with probability $1/2$, and all the clones are grouped into clauses uniformly at random. We study properties of Random SAT problems generated this way for a wide range of degree distributions.

References

- 1 Oleksii Omelchenko, Andrei A. Bulatov. Satisfiability and algorithms for non-uniform random k -SAT. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(5):3886-3894, 2021. URL: <https://ojs.aaai.org/index.php/AAAI/article/view/16507>.
- 2 Oleksii Omelchenko, Andrei A. Bulatov. Satisfiability threshold for power law random 2-SAT in configuration model. *Theoretical Computer Science*, 888:70-94, 2021. doi: 10.1016/j.tcs.2021.07.028.
- 3 Oleksii Omelchenko, Andrei A. Bulatov. Concentration inequalities for sums of random variables, each having power bounded tails. [arXiv:1903.02529](https://arxiv.org/abs/1903.02529), 2019.

3.9 Convergence law for random permutations

Valentin Féray (CNRS – Vandoeuvre-lès-Nancy, FR)

License  Creative Commons BY 4.0 International license
© Valentin Féray

Joint work of Michael H. Albert, Mathilde Bouvel, Valentin Féray, Marc Noy

There are two natural ways to see permutations as models of some logical theory: either as bijections from a set A to itself or as a pair of linear orders on a set A . In this talk we discuss the question of finding convergence laws for models of random permutations. In particular, we prove a convergence law for uniform 231-avoiding permutations, seen as pairs of orders.

3.10 Logic and property testing on graphs of bounded degree

Isolde Adler (University of Leeds, GB)

License  Creative Commons BY 4.0 International license
© Isolde Adler

Joint work of Isolde Adler, Polly Fahey, Frederik Harwath, Noleen Köhler, Pan Peng

Property testing (for a property P) asks for a given graph, whether it has property P , or is “structurally far” from having that property. A “testing algorithm” is a probabilistic algorithm that answers this question with high probability correctly, by only looking at small parts of the input. Testing algorithms are thought of as “extremely efficient”, making them relevant in the context of large data sets.

In this talk I will present recent positive and negative results about testability of properties definable in first-order logic and monadic second-order logic on classes of bounded-degree graphs.

3.11 Limiting probabilities of first order properties of random sparse graphs

Marc Noy (UPC Barcelona Tech, ES)

License  Creative Commons BY 4.0 International license
© Marc Noy

Joint work of Alberto Larrauri, Tobias Müller, Marc Noy

Main reference Lázaro Alberto Larrauri, Tobias Müller, Marc Noy: “Limiting probabilities of first order properties of random sparse graphs and hypergraphs”, *Random Struct. Algorithms*, Vol. 60(3), pp. 506–526, 2022.

URL <https://doi.org/10.1002/rsa.21041>

Let G_n be the binomial random graph $G(n, p = c/n)$ in the sparse regime, which as is well-known undergoes a phase transition at $c = 1$. Lynch [1] showed that for every first order sentence ϕ , the limiting probability that G_n satisfies ϕ as $n \rightarrow \infty$ exists, and moreover it is an analytic function of c . In this paper we consider the closure $\overline{L_c}$ in the interval $[0, 1]$ of the set L_c of all limiting probabilities of first order sentences in G_n . We show that there exists a critical value $c_0 \approx 0.93$ such that $\overline{L_c} = [0, 1]$ when $c \geq c_0$, whereas $\overline{L_c}$ misses at least one subinterval when $c < c_0$.

References

- 1 James F. Lynch. Probabilities of sentences about very sparse random graphs. *Random Structures & Algorithms*, 3(1):33–53, 1992. doi:10.1002/rsa.3240030105.

3.12 Logical limit laws for Mallows random permutations

Tobias Müller (University of Groningen, NL)

License  Creative Commons BY 4.0 International license
© Tobias Müller

Joint work of Tobias Müller, Fiona Skerman, Teun Verstraaten

The Mallows distribution samples a permutation on $1, \dots, n$ at random, where each permutation has probability proportional to $q^{\text{inv}(\pi)}$, where $q > 0$ is a parameter and $\text{inv}(\cdot)$ denotes the number of inversions of π . So in particular, setting $q = 1$ we retrieve the uniform distribution.

In the talk, I discussed some preliminary results on logical limit laws for Mallows random permutations, using two different logical languages of permutations, called “theory of one bijection” (TOOB) and “theory of two orders” (TOTO) by Albert, Bouvel and Feray.

3.13 On first order model checking on Erdős-Rényi graphs

Peter Rossmanith (RWTH Aachen, DE)

License  Creative Commons BY 4.0 International license
© Peter Rossmanith

Clique can be solved in expected FPT time on uniformly distributed graphs of size n while this is not clear for Dominating Set. We show that it is indeed unlikely that Dominating Set can be solved efficiently on random graphs: If yes, then every first-order expressible graph property can be solved in expected FPT time, too. Furthermore, this remains true when we consider random graphs with an arbitrary constant edge probability. There is a very simple problem on random matrices that is equally hard to solve on average: Given a square boolean matrix, are there k rows whose logical AND is the zero vector?

3.14 Probabilistic databases – overview

Dan Suciu (University of Washington – Seattle, US)

License © Creative Commons BY 4.0 International license
© Dan Suciu

Main reference Dan Suciu: “Probabilistic Databases for All”, in Proc. of the 39th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, PODS 2020, Portland, OR, USA, June 14-19, 2020, pp. 19–31, ACM, 2020.

URL <https://doi.org/10.1145/3375395.3389129>

In probabilistic databases the data is uncertain and is modeled by a probability distribution. The central problem in probabilistic databases is query evaluation, which requires performing not only traditional data processing such as joins, projections, unions, but also probabilistic inference in order to compute the probability of each item in the answer. At their core, probabilistic databases are a proposal to integrate logic with probability theory. This talk gives an overview of the probabilistic data models and the main results on query evaluation on probabilistic databases.

3.15 Explaining query answers through probabilistic databases

Benny Kimelfeld (Technion – Haifa, IL)

License © Creative Commons BY 4.0 International license
© Benny Kimelfeld

Joint work of Antoine Amarilli, Leopoldo Bertossi, Daniel Deutch, Nave Frost, Benny Kimelfeld, Ester Livshits, Mikaël Monet

Main reference Daniel Deutch, Nave Frost, Benny Kimelfeld, Mikaël Monet: “Computing the Shapley Value of Facts in Query Answering”, CoRR, Vol. abs/2112.08874, 2021.

URL <https://arxiv.org/abs/2112.08874>

One of the challenges in explanations for data-analysis tools is the quantification of the responsibility of individual data items to the overall result. Nowadays a common approach is to deploy concepts from the theory of cooperative games. A primary example is the Shapley value – a conventional and well-studied function for determining the contribution of a player to the coalition. I will describe our recent research on the computation of the Shapley value (and values of a similar nature) of a database tuple as its contribution to the result of a query. It turns out that the complexity of this task has tight connections to the complexity of query evaluation over tuple-independent probabilistic databases. Moreover, this connection highlights the need to understand the complexity of query evaluation in cases where probability assignments are restricted, such as the uniform case where each tuple has the probability 0.5 (and the task is to count the database subsets that satisfy the query). I will present a recent resolution of the uniform case, as well as every case where the probability is fixed in every relation, for an important class of database queries.

3.16 Probabilistic query evaluation with bag semantics

Peter Lindner (EPFL Lausanne, CH)

License © Creative Commons BY 4.0 International license
© Peter Lindner

Joint work of Martin Grohe, Peter Lindner, Christoph Standke

Main reference Martin Grohe, Peter Lindner, Christoph Standke: “Probabilistic Query Evaluation with Bag Semantics”, CoRR, Vol. abs/2201.11524, 2022.

URL <https://arxiv.org/abs/2201.11524>

Probabilistic databases (PDBs) are usually introduced as probability spaces over the subsets of a finite set of possible facts. Probabilistic query evaluation (PQE) is the problem of finding the probability that a Boolean query evaluates to true on a given PDB. The data complexity of PQE is well-understood for the class of unions of conjunctive queries on tuple-independent PDBs: for every query, the problem can be classified to be either solvable in polynomial time, or is #P-hard [1].

In this talk, we discuss a version of PQE, where both the query evaluation, and the probabilistic databases use a bag semantics. That is, the input PDB may contain duplicates, and the query evaluation takes tuple multiplicities into account. Our main focus lies on self-join free conjunctive queries where we obtain a dichotomy similar to the one above. This is achieved by combining known results for the set version of the problem with novel techniques that are required to handle distributions over multiplicities.

References

- 1 Nilesch Dalvi and Dan Suciu. The dichotomy of probabilistic inference for unions of conjunctive queries. *Journal of the ACM*, 59(6):30:1–30:87, 2012. doi:10.1145/2395116.2395119.

3.17 Limits of graphs

Patrice Ossona de Mendez (EHESS – Paris, FR)

License © Creative Commons BY 4.0 International license
© Patrice Ossona de Mendez

Joint work of Jaroslav Nešetřil, Patrice Ossona de Mendez

How to represent limits of networks? In this survey, several kinds of limits are considered: elementary limits, left limits, local limits, and structural limits, as well as different types of limits objects (distributional or analytical). A particular attention is given to the notion of structural limits, based on the convergence of the satisfaction probabilities of first-order formulas, which generalize classical notions and offers a both a distributional limit objects (as a probability distribution on a Stone space) and an analytic limit object (a modeling, which is a totally Borel structure on a standard probability space).

3.18 Zero-one laws and almost sure valuations of first-order logic in semiring semantics

Matthias Naaf (RWTH Aachen, DE)

License © Creative Commons BY 4.0 International license
© Matthias Naaf

Joint work of Erich Grädel, Hayyan Helal, Matthias Naaf, Richard Wilke

Main reference Erich Grädel, Hayyan Helal, Matthias Naaf, Richard Wilke: “Zero-One Laws and Almost Sure Valuations of First-Order Logic in Semiring Semantics”, CoRR, Vol. abs/2203.03425, 2022.

URL <https://doi.org/10.48550/arXiv.2203.03425>

Semiring semantics evaluates logical statements by values in some commutative semiring K , and random semiring interpretations, induced by a probability distribution on K , generalise random structures. In this talk, we investigate how the classical 0-1 laws of Glebskii et al. and Fagin generalise to semiring semantics.

Using algebraic representations of FO-formulae, we show that for many semirings, including min-max-semirings and the natural semiring, 0-1 laws hold under reasonable assumptions on the probability distribution. We can further partition the FO-sentences into classes $(\Phi_j)_{j \in K}$ for all semiring elements j , such that all sentences in Φ_j almost surely evaluate to j in random semiring interpretations. For finite min-max and lattice semirings, this partition actually collapses to just three classes Φ_0 , Φ_1 and Φ_ε of sentences that, respectively, almost surely evaluate to 0, to the greatest semiring value 1, and to the smallest non-zero value ε . Computing this almost sure valuation is a PSPACE-complete problem.

3.19 Moser-Tardos algorithm with small number of random bits

Oleg Pikhurko (University of Warwick – Coventry, GB)

License © Creative Commons BY 4.0 International license
© Oleg Pikhurko

Joint work of Endre Csóka, Lukasz Grabowski, András Máthé, Oleg Pikhurko, Konstantinos Tyros

Main reference Endre Csóka, Lukasz Grabowski, András Máthé, Oleg Pikhurko, Konstantinos Tyros: “Moser-Tardos Algorithm with small number of random bits”, CoRR, Vol. abs/2203.05888, 2022.

URL <https://doi.org/10.48550/arXiv.2203.05888>

We present a variant of the parallel Moser-Tardos Algorithm [1] and show that, for a class of problems whose dependency graphs have some subexponential growth, the expected number of random bits used by the algorithm is constant. In particular the expected number of used random bits is independent from the total number of variables. This is achieved by using the same random bits to resample variables which are far enough in the dependency graph.

References

- 1 Robin A. Moser and Gábor Tardos. A constructive proof of the general Lovász local lemma. *Journal of the ACM*, 57(2):11:1–11:15, 2010. doi:10.1145/1667053.1667060.

3.20 The repeated insertion model (RIM): probabilistic inference and applications

Batya Kenig (Technion – Haifa, IL)

License © Creative Commons BY 4.0 International license

© Batya Kenig

Joint work of Lovro Ilijasić, Batya Kenig, Benny Kimelfeld, Haoyue Ping, Julia Stoyanovich

Main reference Batya Kenig, Lovro Ilijasic, Haoyue Ping, Benny Kimelfeld, Julia Stoyanovich: “Probabilistic Inference Over Repeated Insertion Models”, in Proc. of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018, pp. 1897–1904, AAAI Press, 2018.

URL <https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/16252>

Distributions over rankings are used to model user preferences in elections, commerce, and more. The Repeated Insertion Model (RIM) gives rise to various known probability distributions over rankings, in particular to the popular Mallows model. However, probabilistic inference over RIM is provably intractable in the general case. In this talk, I will describe an algorithm for computing the marginal probability of an arbitrary partially ordered set over RIM. The complexity of the algorithm is captured in terms of a new measure termed the “cover width”. I will briefly discuss an application of RIM for the task of estimating the probability of an outcome in an election over probabilistic votes.

References

- 1 Batya Kenig and Benny Kimelfeld. Approximate inference of outcomes in probabilistic elections. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01):2061–2068, 2019. doi:10.1609/aaai.v33i01.33012061.

3.21 Repairs, measures, and complexity for constraints violations in databases

Sudeepa Roy (Duke University – Durham, US)

License © Creative Commons BY 4.0 International license

© Sudeepa Roy

Joint work of Benny Kimelfeld, Ester Livshits, Ihab Ilyas, Rina Kochirgan, Sudeepa Roy, Segev Tsur

Noisy or inconsistent databases that violate one or more integrity constraints expected to hold on the database are abundant in practice. First, I will discuss the complexity of computing an optimal “repair” for an inconsistent database where integrity constraints are Functional Dependencies, both in terms of “subset repairs” (by a minimum number of tuple deletion) and “update repairs” (by a minimum number of value or cell updates), and draw a connection to the complexity of the “most probable database” problem. Then I will discuss theoretical properties of “measures” of inconsistencies in a database where integrity constraints are violated, which can be useful for reliability estimation for datasets and progress indication in data repair.

3.22 Ordered graphs of bounded twin-width

Szymon Torunczyk (*University of Warsaw, PL*)

License  Creative Commons BY 4.0 International license
© Szymon Torunczyk

We establish a list of characterizations of bounded twin-width for hereditary classes of totally ordered graphs: as classes of at most exponential growth studied in enumerative combinatorics, as NIP classes studied in model theory, as classes that do not transduce the class of all graphs studied in finite model theory, and as classes for which model checking first-order logic is fixed-parameter tractable studied in algorithmic graph theory.

This has several consequences. First, it allows us to show that every hereditary class of ordered graphs either has at most exponential growth, or has at least factorial growth. This settles a question first asked by Balogh, Bollobás, and Morris [1] on the growth of hereditary classes of ordered graphs, generalizing the Stanley-Wilf conjecture/Marcus-Tardos theorem. Second, it gives a fixed-parameter approximation algorithm for twin-width on ordered graphs. Third, it yields a full classification of fixed-parameter tractable first-order model checking on hereditary classes of ordered binary structures. Finally, it provides a model-theoretic characterization of classes with bounded twin-width.

References

- 1 József Balogh, Béla Bollobás, and Robert Morris. Hereditary properties of partitions, ordered graphs and ordered hypergraphs. *European Journal of Combinatorics*, 27(8):1263–1281, 2006. doi:10.1016/j.ejc.2006.05.004.

Participants

- Aida Abiad
TU Eindhoven, NL
- Isolde Adler
University of Leeds, GB
- Albert Atserias
UPC Barcelona Tech, ES
- Manuel Bodirsky
TU Dresden, DE
- Andrei A. Bulatov
Simon Fraser University –
Burnaby, CA
- Anuj Dawar
University of Cambridge, GB
- Valentin Féray
CNRS – Vandoeuvre-lès-
Nancy, FR
- Erich Grädel
RWTH Aachen, DE
- Benny Kimelfeld
Technion – Haifa, IL
- Peter Lindner
EPFL Lausanne, CH
- Tobias Müller
University of Groningen, NL
- Matthias Naaf
RWTH Aachen, DE
- Marc Noy
UPC Barcelona Tech, ES
- Oleg Pikhurko
University of Warwick –
Coventry, GB
- Peter Rossmanith
RWTH Aachen, DE
- Dan Suciu
University of Washington –
Seattle, US
- Lidia Tendra
University of Opole, PL
- Teun Verstraaten
University of Groningen, NL



Remote Participants

- Andreas R. Blass
University of Michigan –
Ann Arbor, US
- Valentin Goranko
University of Stockholm, SE
- Batya Kenig
Technion – Haifa, IL
- Lefteris M. Kirousis
University of Athens, GR
- Phokion G. Kolaitis
University of California –
Santa Cruz, US
- Daniel Král’
Masaryk University – Brno, CZ
- Jerzy Marcinkowski
University of Wrocław, PL
- Patrice Ossona de Mendez
EHESS – Paris, FR
- Sudeepa Roy
Duke University – Durham, US
- Fiona Skerman
Uppsala University, SE
- Christoph Standke
RWTH Aachen, DE
- Szymon Torunczyk
University of Warsaw, PL
- Maksim Zhukovskii
MIPT – Dolgoprudny, RU

Computation and Reconfiguration in Low-Dimensional Topological Spaces

Maike Buchin^{*1}, Anna Lubiw^{*2}, Arnaud de Mesmay^{*3},
Saul Schleimer^{*4}, and Florestan Brunck^{†5}

- 1 Faculty of Computer Science, Ruhr-Universität Bochum, DE.
maike.buchin@rub.de
- 2 Cheriton School of Computer Science, University of Waterloo, CA.
alubiw@waterloo.ca
- 3 LIGM, CNRS, Univ. Gustave Eiffel, ESIEE Paris, Marne-la-Vallée, FR.
arnaud.de-mesmay@univ-eiffel.fr
- 4 Mathematics Institute, University of Warwick, Coventry, GB.
s.schleimer@warwick.ac.uk
- 5 IST Austria – Klosterneuburg, AT. florestan.brunck@ist.ac.at

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 22062 “Computation and Reconfiguration in Low-Dimensional Topological Spaces”. The seminar consisted of a small collection of introductory talks, an open problem session, and then the seminar participants worked in small groups on problems on reconfiguration within the context of objects as diverse as elimination trees, morphings, curves on surfaces, translation surfaces and Delaunay triangulations.

Seminar February 6–11, 2022 – <http://www.dagstuhl.de/22062>

2012 ACM Subject Classification Human-centered computing → Graph drawings; Theory of computation → Computational geometry; Mathematics of computing → Geometric topology

Keywords and phrases Geometric Topology, Computational Geometry, Graph Drawing, Reconfiguration, Dagstuhl Seminar

Digital Object Identifier 10.4230/DagRep.12.2.17

1 Executive Summary

Maike Buchin (Faculty of Computer Science, Ruhr-Universität Bochum, maike.buchin@rub.de)

Anna Lubiw (Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, Canada, alubiw@waterloo.ca)

Arnaud de Mesmay (LIGM, CNRS, Univ. Gustave Eiffel, ESIEE Paris, F-77454 Marne-la-Vallée, France, arnaud.de-mesmay@univ-eiffel.fr)

Saul Schleimer (Mathematics Institute, University of Warwick, Coventry CV4 7AL, United Kingdom, s.schleimer@warwick.ac.uk)

License  Creative Commons BY 4.0 International license
© Maike Buchin, Anna Lubiw, Arnaud de Mesmay and Saul Schleimer

This seminar was proposed as a followup to the Dagstuhl Seminars 17072: “Applications of Topology to the Analysis of 1-Dimensional Objects” and 19352: “Computation in Low-Dimensional Geometry and Topology”. The goal of these seminars was to bring together

* Editor / Organizer

† Editorial Assistant / Collector



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

Computation and Reconfiguration in Low-Dimensional Topological Spaces, *Dagstuhl Reports*, Vol. 12, Issue 2, pp. 17–66

Editors: Maike Buchin, Anna Lubiw, Arnaud de Mesmay, Saul Schleimer, and Florestan Brunck



DAGSTUHL
REPORTS Dagstuhl Reports

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

researchers from different communities who are working on low-dimensional topological spaces (curves, embedded graphs, knots, surfaces, three-manifolds), in order to foster collaborations and synergies. Indeed, while the mathematical study of these objects has a rich and old history, the study of their algorithmic properties is still in its infancy, and new questions and problems keep coming from theoretical computer science and more applied fields, yielding a fresh and renewed perspective on computation in topological spaces.

The success of previous seminars demonstrated that research in low-dimensional topology is very active and fruitful, and also that there was a strong demand for a new seminar gathering researchers from the various involved communities, namely geometric topology and knot theory, computational geometry and topology, all the way to graph drawing and trajectory analysis.

For this iteration we placed a particular emphasis on topics related to geometric and topological **reconfiguration**: How can one structure be changed into another? How far apart are two structures? Such questions lie at the heart of various geometric problems such as computing the Fréchet distance as a way to quantify curve similarity, or morphing between two versions of a common graph. In many cases, the combinatorics and the geometry of a reconfiguration space also emerged as important objects of study: examples include associahedra, the flip graphs of triangulations, and the curve complexes in geometric topology.

The seminar started with four overview talks given by researchers in geometric topology, computational geometry, topological dynamics, and graph drawing to motivate and propose open problems that would fit the diverse backgrounds of participants and the specific focus on reconfiguration chosen for this year's workshop. This was followed by an open problem session where we gathered fifteen open problems, some of which were circulated in advance of the meeting. The remainder of the week was spent actively working on solving these problems in small groups.

The Covid pandemic prevented many participants from attending the seminar physically, and the entirety of the seminar took place in a hybrid setting, with most working groups featuring both online and physical participants. In order to coordinate the progress, we used Coauthor, a tool designed for by Erik Demaine (MIT), which greatly facilitated the collaborations, and also allowed participants to have a record of the work when the seminar concluded. We also held two daily progress report meetings, allowing people to share progress and allow people to switch groups. In addition to the traditional hike, a virtual social meeting was held on Gather.town to foster interactions between the online and the physical participants.

We now briefly describe the problems that have been worked on, with a more in-depth survey of the problems and the progress being done being featured farther down in this Dagstuhl Report. Some more open problems that have been proposed but not worked on are also listed at the end of the document.

Two groups worked on questions pertaining to reconfiguring curves in the plane and on surfaces. The group 4.1 investigated problems inspired by nonograms, where one aims at introducing switches at intersections of curves in the plane to remove so-called popular faces. The group 4.5 looked at the reconfiguration graph obtained under the action of local moves on minimal closed (multi-)curves on surfaces, and whether such multi-curves could be realized as the set of geodesics of some hyperbolic metric on the surface.

A different flavor of surfaces was studied by the group 4.4, who investigated how square-tiled surfaces could be transformed under the action of shears of cylinder blocks.

The working group 4.2 studied the longstanding problem of the computational complexity of evaluation the rotation distance between elimination trees in graphs. A different flip graph, namely the one of order- k Delaunay triangulations was the topic of study of group 4.7.

Finally, two groups worked on motion of discrete objects in different contexts. The group 4.3 initiated a generalization of the classical theory of morphings of planar graph when one allows the morph to go through a third dimension. The group 4.6 investigated Turning machines, which is a simple model of molecular robot aiming to fold into specific shapes.

All in all, the seminar fostered a highly collaborative research environment by allowing researchers from very diverse backgrounds to work together on precise problems. While the hybrid setting proved to be a significant challenge, the quality of the equipment at Dagstuhl and the online tools that were used provided a practical way for all the participants to interact and to make progress on problems related to reconfiguration in geometric and topological settings.

2 Table of Contents

Executive Summary

<i>Maike Buchin, Anna Lubiw, Arnaud de Mesmay and Saul Schleimer</i>	17
--	----

Overview of Talks

Playing Puzzles on Square-Tiled Surfaces <i>Hugo Parlier</i>	22
Flip Graphs and Polytopes <i>Jean Cardinal</i>	22
Geometry of Large Genus Flat Surfaces and Open Problems on Square-Tiled Surfaces <i>Élise Goujard</i>	23
Morphing Graph Drawings <i>Fabrizio Fratti</i>	23

Working Groups

Reconfiguring Popular Faces <i>Florestan Brunck, Hsien-Chih Chang, Maarten Löffler, Tim Ophelders, Lena Schlipf</i>	24
Rotation Distance between Elimination Tree <i>Maike Buchin, Jean Cardinal, Linda Kleist, Boris Klemz, Anna Lubiw, Torsten Mütze, Alexander Neuhaus, Lionel Pournin</i>	35
Morphing Planar Graph Drawings Through 3D <i>Kevin Buchin, Will Evans, Fabrizio Frati, Irina Kostitsyna, Maarten Löffler, Tim Ophelders, Carola Wenk, Alexander Wolff</i>	40
Statistics of Square Tiled Surfaces and the shearing block game <i>Vincent Delecroix, Elise Goujard, Luke Jeffreys, Hugo Parlier, Saul Schleimer</i> . . .	42
The graph of minimal configurations of a multicurve <i>Francis Lazarus, Mark Bell, Benjamin Burton, Erin Moriarty Wolf Chambers, Vincent Delecroix, Niloufar Fuladi, Luke Jeffreys, Francis Lazarus, Arnaud de Mesmay, Hugo Parlier, Saul Schleimer, Éric Colin de Verdière, Hsien-Chih Chang, Jean Cardinal</i>	49
Turning Machines <i>Kevin Buchin, Jean Cardinal, Will Evans, Irina Kostitsyna, Alexander Wolff</i> . . .	56
Flips in Higher Order Delaunay triangulations <i>Rodrigo Silveira, Anne Driemel, Jonathan Spreer, Lena Schlipf</i>	59

Open Problems

Representing Graphs by Polygon Contact in 3D <i>Will Evans and Alexander Wolff</i>	60
Diameter of rectangulation flip graph <i>Torsten Mütze</i>	61
Balanced triangulation Gray codes <i>Torsten Mütze</i>	63

Pattern-Matching on Braids <i>Tim Ophelders</i>	64
Participants	65
Remote Participants	65

3 Overview of Talks

3.1 Playing Puzzles on Square-Tiled Surfaces

Hugo Parlier (*University of Luxembourg – Esch-sur-Alzette, Luxembourg, hugo.parlier@uni.lu*)

License  Creative Commons BY 4.0 International license
© Hugo Parlier

Joint work of Hugo Parlier, Paul Turner, Mario Gutierrez, Reyna Juarez, Mark Bell, Lionel Pournin

This talk will be about a project aiming to illustrate geometry through puzzles. The puzzles are played on (square-tiled) surfaces, and have natural configuration graphs with a geometry of their own. These graphs are reminiscent of combinatorial graphs used in the study of moduli spaces of surfaces which can be visualized in similar ways.

The puzzles were created together with Paul Turner, and brought to life together with Mario Gutierrez and Reyna Juarez. The pictures of moduli spaces were created with Mark Bell and Lionel Pournin.

References

- 1 Parlier, Hugo; Pournin, Lionel, Modular flip-graphs of one-holed surfaces. *Eur. J. Comb.* 67, 158-173 (2018).

3.2 Flip Graphs and Polytopes

Jean Cardinal (*Université Libre de Bruxelles – Bruxelles, Belgique, cardinaljean@gmail.com*)

License  Creative Commons BY 4.0 International license
© Jean Cardinal

We present various families of flip graphs that are skeletons of polytopes. We begin with the ubiquitous Associahedron, whose skeleton is the flip graph of triangulations of a convex n -gon and the rotation graph of binary trees with $n - 1$ leaves. We then introduce graph associahedra, an elegant generalization of associahedra whose skeleton is the rotation graph of so-called elimination trees on a given simple connected graph G . Next, we review well-known properties of the flip graph of acyclic orientations of a graph G , the skeleton of the graphical zonotope of G . We then proceed to show that graphical zonotopes and graph associahedra have a common generalization called hypergraphic polytopes, whose skeletons are flip graphs of acyclic orientations of a given hypergraph.

For each family of flip graphs, we mention old and new results on flip distance and Hamiltonicity properties, emphasizing the computational aspects: How hard is it to compute the flip distance between two given objects? Does there exist an efficient Gray code for listing these objects, one flip at a time?

References

- 1 Alexander Postnikov: Permutohedra, associahedra, and beyond. *International Mathematics Research Notices*, 6:1026–1106, 2009.
- 2 Thibault Manneville and Vincent Pilaud: graph properties of graph associahedra. *Séminaire Lotharingien de Combinatoire* 73 (2015), Article B73d
- 3 Marcelo Aguiar and Federico Ardila: Hopf monoids and generalized permutahedra, arXiv:1709.07504, 2017.
- 4 Carolina Benedetti, Nantel Bergeron, John Machacek: Hypergraphic polytopes: combinatorial properties and antipode. arXiv:1712.08848, 2017.

- 5 Jean Cardinal, Stefan Langerman, Pablo Pérez-Lantero: On the diameter of tree associahedra. *Electron. J. Comb.* 25(4): P4.18, 2018.
- 6 Prosenjit Bose, Jean Cardinal, John Iacono, Grigorios Koumoutsos, Stefan Langerman: Competitive online search trees on trees. *SODA 2020*: 1878-1891.
- 7 Jean Cardinal, Lionel Pournin, Mario Valencia-Pabon: Bounds on the diameter of graph associahedra. *LAGOS 2021*: 239-247. (arXiv:2106.16130).
- 8 Jean Cardinal, Arturo I. Merino, Torsten Mütze: Combinatorial generation via permutation languages. IV. Elimination trees. *CoRR abs/2106.16204* (2021). Extended abstract in *SODA 2022*.

3.3 Geometry of Large Genus Flat Surfaces and Open Problems on Square-Tiled Surfaces

Élise Goujard (*University of Bordeaux – Bordeaux, France, elise.goujard@u-bordeaux.fr*)

License  Creative Commons BY 4.0 International license

© Élise Goujard

Joint work of Delecroix, Vincent; Zograf, Peter; Zorich, Anton

Main reference Vincent Delecroix, Elise Goujard, Peter Zograf, Anton Zorich: “Large genus asymptotic geometry of random square-tiled surfaces and of random multicurves”, arXiv, CoRR, Vol. abs. 2007.04740, 2020.

URL <http://dx.doi.org/10.48550/ARXIV.2007.04740>

In this talk we present some results (joint work with Vincent Delecroix, Peter Zograf and Anton Zorich) on the geometry of large genus surfaces and related problems on square-tiled surfaces. The study of the $SL(2, \mathbb{R})$ -dynamics and the geometry of the moduli space of translation surfaces allows to prove equidistribution of square-tiled surfaces with fixed combinatorics in the strata and uncorrelation between horizontal and vertical combinatorics [2], as well as large genus asymptotics for the distribution of cylinders for instance [1]. Some results about the statistics of random square-tiled surfaces with no constraints on the singularities (especially for half-translation square-tiled surfaces) are still open.

References

- 1 V. Delecroix, é. Goujard, P. Zograf, A. Zorich, *Large genus asymptotic geometry of random square-tiled surfaces and of random multicurves*, arXiv:2007.04740.
- 2 V. Delecroix, é. Goujard, P. Zograf, A. Zorich, *Enumeration of meanders and Masur-Veech volumes*. *Forum Math. Pi* 8 (2020), e4, 80 pp.

3.4 Morphing Graph Drawings

Fabrizio Fratti (*Roma Tre University – Rome, Italy, fratti@dia.uniroma3.it*)

License  Creative Commons BY 4.0 International license

© Fabrizio Fratti

A morph between two geometric shapes is a continuous transformation of one shape into the other. Morphs are useful in many areas of computer science, including Computer Graphics, Animation, and Modeling. This talk surveys known results and algorithms on morphing graph drawings.

The first part of the talk is devoted to morphing algorithms for planar straight-line drawings. Contraction-based algorithms [1, 2], coefficient-interpolation-based algorithms [4], and one-coefficient-at-a-time algorithms [7] are described. The running time and the resolution of the described morphing algorithms are discussed.

The second part of the talk deals with other graph drawing styles. In particular, morphs between non-planar graph drawings [3], three-dimensional morphs [5], and upward morphs [6] are discussed

References

- 1 S.S. Cairns, *Deformations of plane rectilinear complexes*, “The American Mathematical Monthly”, volume 51, 5, 247–252, Mathematical Association of America, 1944.
- 2 Soroush Alamdari and Patrizio Angelini and Fidel Barrera-Cruz and Timothy M. Chan and Giordano Da Lozzo and Giuseppe Di Battista and Fabrizio Frati and Penny Haxell and Anna Lubiw and Maurizio Patrignani and Vincenzo Roselli and Sahil Singla and Bryan T. Wilkinson, *How to Morph Planar Graph Drawings*, SIAM J. Comput., volume 46, 2, 824–852, 2017.
- 3 Patrizio Angelini and Michael A. Bekos and Fabrizio Montecchiani and Maximilian Pfister, *On Morphing 1-Planar Drawings*, 47th International Workshop on Graph-Theoretic Concepts in Computer Science (WG), Lecture Notes in Computer Science, 12911, 270–282, Springer, 2021.
- 4 Michael S. Floater and Craig Gotsman, *How to morph tilings injectively*, Journal of Computational and Applied Mathematics, volume 101, 1, 117–129, 1999, doi:10.1016/S0377-0427(98)00202-7.
- 5 Elena Arseneva and Prosenjit Bose and Pilar Cano and Anthony D’Angelo and Vida Dujmovic and Fabrizio Frati and Stefan Langerman and Alessandra Tappini, *Pole Dancing: 3D Morphs for Tree Drawings*, J. Graph Algorithms Appl., volume 23, 3, 579–602, 2019, <https://doi.org/10.7155/jgaa.00503>.
- 6 Giordano Da Lozzo and Giuseppe Di Battista and Fabrizio Frati and Maurizio Patrignani and Vincenzo Roselli, *Upward Planar Morphs*, Algorithmica, volume 82, 10, 2985–3017, 2020, <https://doi.org/10.1007/s00453-020-00714-6>, doi:10.1007/s00453-020-00714-6.
- 7 Jeff Erickson and Patrick Lin, *Planar and Toroidal Morphs Made Easier*, 29th International Symposium on Graph Drawing and Network Visualization (GD), Lecture Notes in Computer Science, Springer, 2021.

4 Working Groups

4.1 Reconfiguring Popular Faces

Florestan Brunck (Institute of Science and Technology, Austria, florestan.brunck@ist.ac.at)

Hsien-Chih Chang (Dartmouth College, USA, hsien-chih.chang@dartmouth.edu)

Maarten Löffler (Utrecht University, the Netherlands, m.loffler@uu.nl)

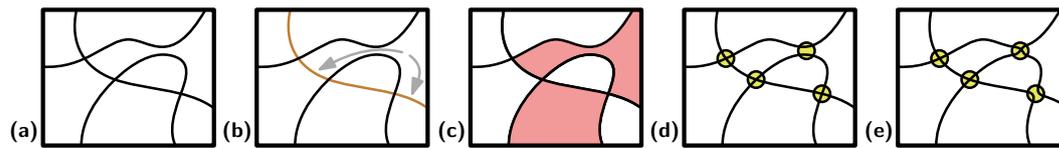
Tim Ophelders (Utrecht University & TU Eindhoven, the Netherlands, t.a.e.ophelders@uu.nl)

Lena Schlipf (Universität Tübingen, Germany, lena.schlipf@uni-tuebingen.de)

License  Creative Commons BY 4.0 International license

© Florestan Brunck, Hsien-Chih Chang, Maarten Löffler, Tim Ophelders, Lena Schlipf

Let \mathcal{A} be a set of curves which lie inside the area bounded by a closed curve F , called the *frame*. All curves in \mathcal{A} are either *closed* or they are *open* with a start and end point on F . We refer to \mathcal{A} as a *curve arrangement*, see Figure 1a. We consider only *simple* arrangements, where no three curves meet in a point, and all intersections are transversal crossings (no tangencies). The arrangement \mathcal{A} can be seen as an embedded multigraph whose vertices are crossings between curves and whose edges are *curve segments*. \mathcal{A} subdivides the region bounded by F into *faces*. We call a face *popular* when it is incident to multiple curve segments belonging to the same curve in \mathcal{A} (see Figures 1b-c).



■ **Figure 1** (a) An arrangement of curves inside a frame. (b) The red curve is incident to the top right face in two disconnected segments, making the face *popular*. (c) All popular faces are highlighted. (d) A set of switches. (e) A possible reconfiguration after which no more faces are popular.

► **Proposition 1.** *We do not like popular faces.*

Now, let a *switch* be a local area in which we are allowed to reroute the curves of \mathcal{A} (see Figures 1d-e).

► **Problem 4.1.** *Given a curve arrangement and a set of switches, can we reconfigure the curves so as to remove all, or as many as possible, popular faces? And can we minimize the number of switch operations?*

Once we are given a set of switches, the question above is essentially combinatorial. Now suppose the switches are not given in advance, but we allow curves to be reconfigured whenever they are “sufficiently close” to each other.

► **Observation 1.** *If all intersections are switches, then it is possible to remove all popular faces.*

Proof. Simply set every switch to any non-intersecting state. Then each curve bounds a single face on each side, and thus no face is popular. ◀

► **Problem 4.2.** *Is there a reasonable way to geometrically determine a set of switches? And how does this influence the complexity of Problem 4.1?*

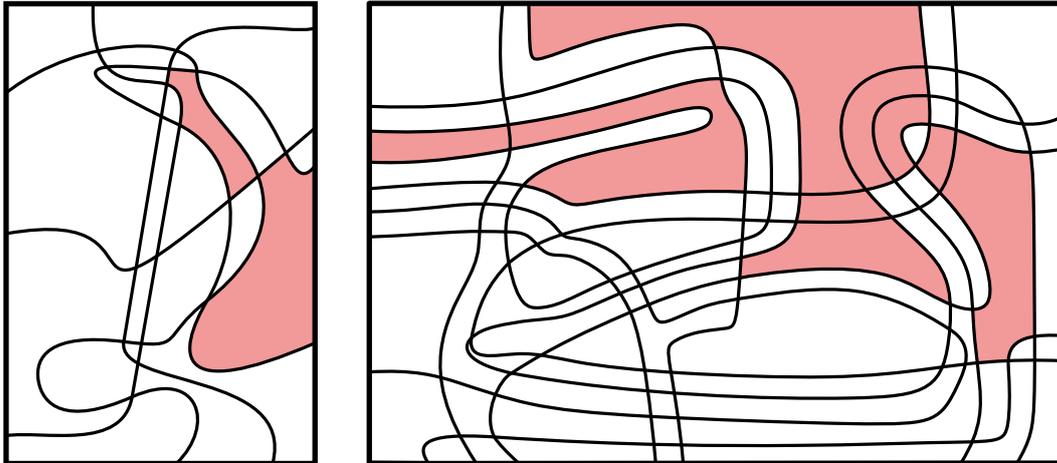
Finally, suppose that our curves have *fixed* parts and *flexible* parts. That is, each curve is a smooth concatenation of pieces, each of which is either fixed or flexible. Fixed pieces may never be altered. Flexible pieces may be changed.

► **Problem 4.3.** *Given a curve arrangement with fixed and flexible pieces, can we reconfigure the flexible parts of the curves so as to remove all popular faces? Can we stay as close to the original arrangement as possible?*

Motivation & Background

Our question is motivated by the problem of generating *curved nonograms*. Nonograms, also known as *Japanese puzzles*, *paint-by-numbers*, or *griddlers*, are a popular puzzle type where one is given an empty grid and a set of *clues* on which grid cells need to be colored. A clue consists of a sequence of numbers specifying the numbers of consecutive filled cells in a row or column. A solved nonogram typically results in a picture (see Figure 2 (a)). There is quite some work in the literature on the difficulty of solving nonograms [1].

Van de Kerkhof et al. introduced *curved nonograms*, a variant in which the puzzle is no longer played on a grid but on any arrangement of curves [4] (see Figure 2b). In curved nonograms, a clue specifies the numbers of filled faces of the arrangement in the sequence of faces that are incident to a common curve on one side. Van de Kerkhof et al. focus on



■ **Figure 4** Some examples of real puzzles (without the clues) with all popular faces highlighted.

One possibility to generate nonograms of a specific complexity would be to take an existing generator and modify the output. Recently, de Nooijer [2] investigated how this might be done by *inserting* a new curve into the arrangement; see also [3]. Problems 1-3 explore a different approach, geared towards the generation of *basic* puzzles. Of course, another interesting question is what can be done to generate *advanced* puzzles.

► **Problem 4.4.** *How do Problems 1-3 change when the goal is only to remove all self-intersections?*

Results

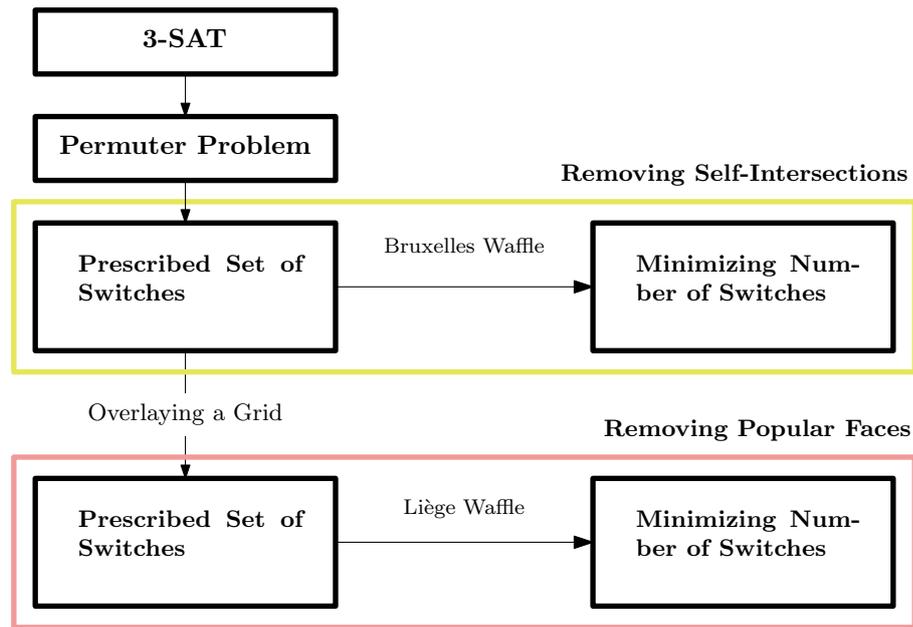
In this section, we present NP-hardness results and bring a negative answer to both Problem 4.1 and its self-intersection variant (Problem 4.4). We begin by examining the self-intersection problem and derive Theorem 1, from which we later reduce our original question about the removal of popular faces, as well as subsequent extensions concerning the minimal number of switch operations required. The organisation of our results is summarized in Fig. 5.

Base Results.

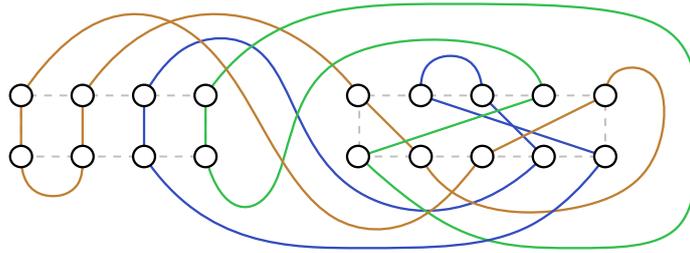
► **Theorem 1.** *Given a curve arrangement and a prescribed set of switches, it is NP-hard to decide whether it is possible to configure the switches such that the resulting arrangement has no self-intersections.*

The proof of Theorem 1 is ultimately a reduction from 3-SAT, but involves an intermediate reduction through a naturally related problem, which we call the *permuter problem*.

The Permuter Problem and its Reduction from 3-SAT. Consider the straight-line-edges drawing of the complete bipartite graph $K_{k,k}$ in the plane, with bi-partition I and O (referred as its *inputs* and *outputs*) such that I and O are evenly distributed on opposite sides of a rectangle R . Fixing the same linear order on the vertices of I and O , the k -*permuter* $\Pi_{k,\sigma}$ is the matching of $K_{k,k}$ associated to the permutation σ of $[n]$. An instance of the permuter problem consists of a finite collection $\{\Pi_{k_i,\sigma_i}\}_{i \in [n]}$ of k_i -permuters, realised in the plane using an associated collection $\{R_i\}_{i \in [n]}$ of rectangles, together with a finite collection of paths $\{P_i\}_{i \in [m]}$, such that:



■ **Figure 5** The flow of successive reductions underlying our NP-hardness proofs.

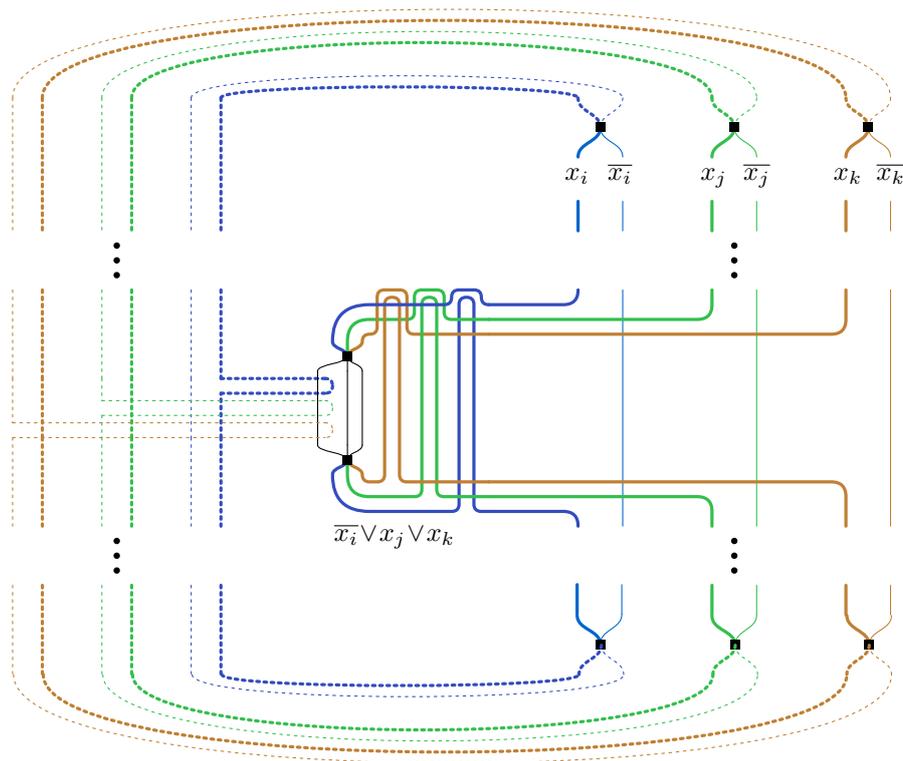


■ **Figure 6** An instance of the k -permuter problem with a 4-permuter and a 5-permuter, assigned with the permutations $\sigma_1 = \text{Id}$ and $\sigma_2 = (12534)$. There are 3 resulting closed loops: one simple (green) and two self-intersecting (blue and brown).

- Every element of $\{P_i\}_{i \in [m]}$ has both its endpoints in one of the k -permuters (possibly the same, and possibly both inputs/outputs).
- Every input and output of every permuter is connected to a unique element of $\{P_i\}_{i \in [m]}$.
- For all $i \in [m]$, for all $j \in [n]$, $P_i \cap R_j = \emptyset$.

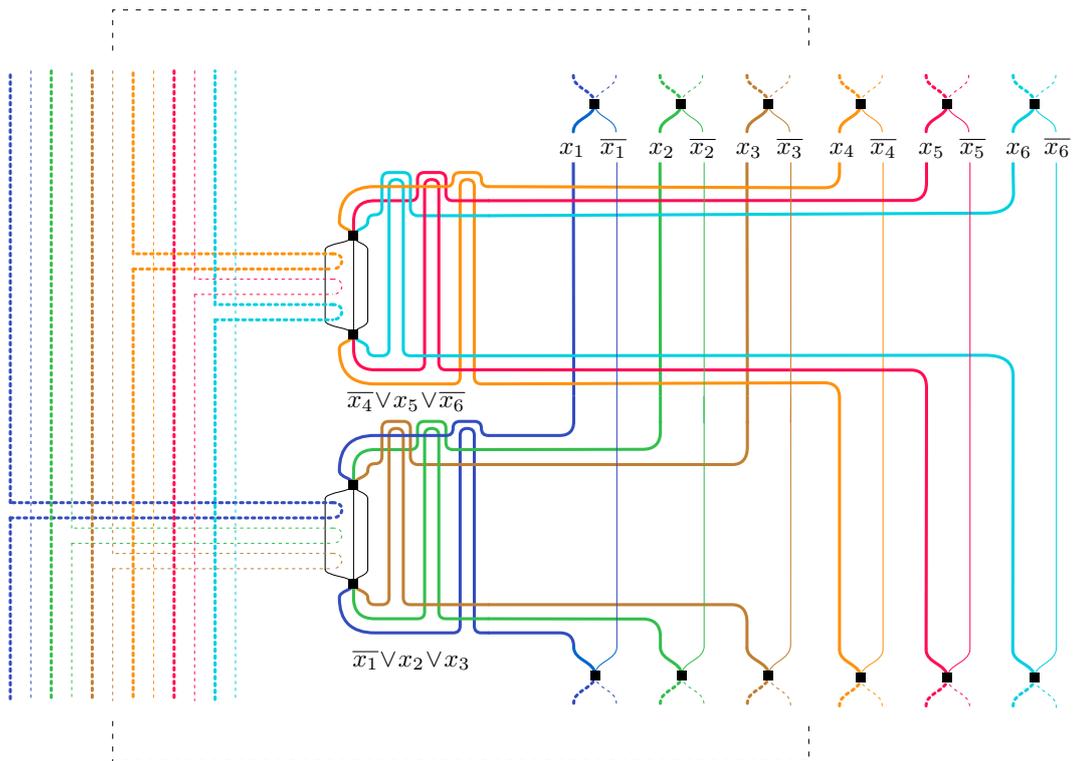
By construction, every path of $\{P_i\}_{i \in [m]}$ belongs to a unique closed loop. The *Permuter problem* then consists in deciding whether or not there exists a choice of n permutations $\sigma_1, \sigma_2, \dots, \sigma_n$ of $[k_1], [k_2], \dots, [k_n]$ such that each resulting loop is *simple* (see Fig. 6).

► **Claim 1.** *The Permuter problem is NP-hard.*



■ **Figure 7** A single 3-clause in the reduction from 3-SAT to the Permuter problem.

Proof. The proof is by reduction from 3-SAT. We shall use two types of gadgets: 2-permuters for variable assignment and 3-permuters for clause verification (see Fig. 7). For simplicity, each k -permuter is depicted by a black box on the diagram, where the value of k is made clear by the number of incoming/outgoing paths. Each different colour in the figure indicates a different variable. The thick or thin dashed lines on the top, bottom and middle-left part of the diagram indicate respectively the false and true literals of each variable. While the thick and thin solid lines in the middle-right section of the diagram indicate respectively the true or false assignment of each variable. Given a Boolean formula with n variables, we construct $2n$ non-crossing semi-circular arcs. We replicate this construction twice to form the top and bottom parts of the diagram. In the middle, we show a single clause gadget, involving two 3-permuters. To simulate the two logical OR of the clause, we proceed as follows: if the corresponding 3-clause involves the variables x_i, x_j and x_k , we select the wire corresponding to their desired truth value in the clause (i.e. thick for x_i , thin for \bar{x}_i) and “drag” them towards the gadget to intersect the same single path chosen among the three paths linking the output of the first 3-permuter to the inputs of the second. By construction, there does not exist a valid permutation assignment to the two 3-permuters which avoids all possible self-intersections with the three black paths if and only if x_i, x_j and x_k all have the wrong truth assignment. Furthermore, to the right of the 3-clause gadget, we have weaved the incoming paths of the first and the outgoing paths of the second in such a way that, if the composition of the two 3-permuters were not the identity, at least one of the resulting closed loop would self-intersect. Thus each such pair of 3-permuters cannot “cheat” and has to compose to the identity. As a consequence, for each of the variables involved, the composition of its two 2-permuters must also be the identity. By construction, there are then exactly two ways of ensuring this is the case: either both 2-permuters are the identity



■ **Figure 8** The instance of the Permuter problem created for the Boolean formula $(\bar{x}_1 \vee x_2 \vee x_3) \wedge (\bar{x}_4 \vee x_5 \vee \bar{x}_6)$.

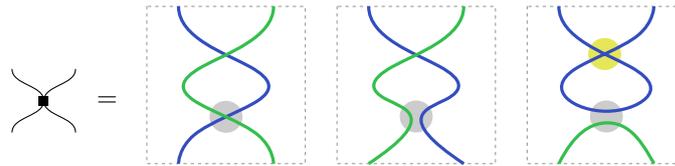
themselves (setting the variable to be true), or both of them correspond to the transposition (12) swapping the inputs (setting the variable to be false). Fig. 8 shows the instance created for the Boolean formula $(\bar{x}_1 \vee x_2 \vee x_3) \wedge (\bar{x}_4 \vee x_5 \vee \bar{x}_6)$. ◀

The Permuter Problem Reduces to the Self-Intersection Problem.

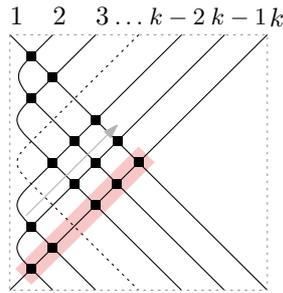
► **Claim 2.** *The problem of configuring a given set of switches to avoid self-intersections reduces from the Permuter problem.*

Proof. The proof of the claim proceeds as follows: we first construct self-intersection gadgets to simulate 2-permuters, take note of the fact that S_n is generated by transpositions and show how to use 2-permuters to construct general k -permuters. The construction for 2-permuters is presented on Fig. 9: the inputs and outputs are connected by two curves weaved into a double coil structure with two intersections, one of which is a switch. The 3 resulting configurations are shown on the figure; only the leftmost two are free of self-intersections.

To simulate a general k -permuter, we introduce the gadget described on Fig. 10. We begin with a k by k square and evenly distribute k inputs and outputs on its top and bottom edges, respectively. For all $i \in [n]$, the top i -th input is connected to the bottom $(n - i)$ -th output by the path of slope -1 which gets reflected into a path of slope 1 upon meeting the left edge of the square. We then insert a total of $\frac{k(k-1)}{2}$ 2-permuter: one at every site where two paths intersect inside the square. While many configurations of this gadget are redundant and yield the same permutation, a short inductive argument shows that it is indeed able to simulate any permutation on k elements. The base case is simply



■ **Figure 9** Constructing a 2-permuter using two “doubly-coiled” curves. The grey disk indicates the only switch available, the yellow disk highlights the self-intersection forbidding the third possibility given by the switch.



■ **Figure 10** Constructing a k -permuter, using $\frac{k(k-1)}{2}$ 2-permuters.

the 2-permuter we previously described. Assume then by induction that the version of our gadget with $k - 1$ inputs and outputs can successfully simulate all permutations of $[k - 1]$. Note that any permutation σ of $[k]$ can be written as the composition of a permutation of $[k - 1]$ followed by the insertion of the element k into one of the k positions before or after one of the $k - 1$ permuted elements. It is thus enough to show that the addition of the last “row” of k 2-permuters (highlighted in pink) can simulate this last insertion step. Labelling the $(k - 1)$ 2-permuters of the $(k - 1)$ -th row according to the direction indicated on Fig. 10, we insert the element k in position i (before the i -th element) by setting all the 2-permuters from positions i to k to swap their inputs, and the remaining i 2-permuters to the identity. This effectively shifts all the elements with positions greater than i by 1 (the permuters reroute their corresponding wires to the segment of slope 1 instead of -1) as we sequentially shift the element k to the left i times (the permuters successively let the k -th input path “slide” on the last path of slope -1). ◀

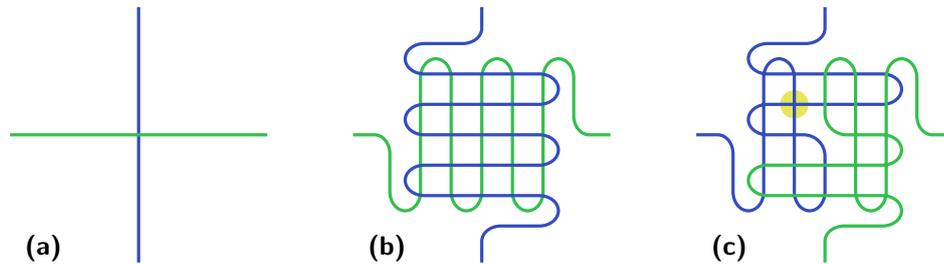
Extensions

We now extend the result from Theorem 1 in several ways, to prove that the problem remains hard when we wish to minimize the number of switch operations, or when the goal is to remove all popular faces rather than self-intersections.

The idea is always to locally alter the reduction in a way that does not affect its global properties.

Minimizing the number of switches. In the previous section, we had a prescribed set of switches; indeed, this is necessary since if we are allowed to switch everywhere, then we can always remove all self-intersections by Observation 1.

However, now consider the scenario in which we wish to minimize the number of switch operations, or, in the decision version, we wish to test for a given k whether there exists a sequence of k switch operations such that the resulting arrangement has no self-intersections. In this scenario, we may or may not have a prescribed set of switches.



■ **Figure 11** The Bruxelles waffle. (a) A crossing between two different curves. (b) We adapt both of the curves so they intersect c^2 times. (c) Any sequence of local switches to emulate a global switch must be of length at least c or lead to a self-intersection.

The idea is to emulate the construction from Section 4.1, but to replace every self-intersection in the construction which is *not* a switch by a *waffle gadget*. Such a gadget is built in such a way that even if *every* intersection in the gadget is a switch, the number of switches required to change its global state is more than a parameter c . If we then choose $c > k$, the result follows since essentially we are never allowed to switch these gadgets.

The Bruxelles Waffle

In this section, we describe the *Bruxelles waffle* (in contrast to the *Liège waffle* which we describe in Section 4.1). The construction is illustrated in Figure 11.

Here we describe the construction in words.

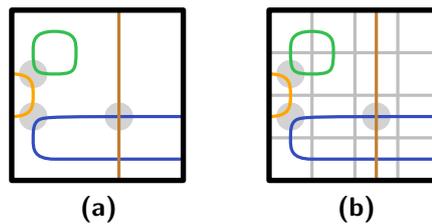
► **Lemma 2.** *In a Bruxelles waffle, any sequence of fewer than c switches must result in an arrangement with either the same combinatorial structure as the original, or at least one self-intersection.*

Proof. Assume that after fewer than c switches, no self-intersections remain. We need to show that opposite terminals lie on the same strand. Assume for a contradiction that opposite terminals do not lie on the same strand. Then the left terminal lies on the same strand as either the top or the bottom terminal. Without loss of generality (by rotational symmetry of the gadget) assume that the left terminal lies on the same strand as the bottom terminal.

Because there are c rows, at least one row has not undergone any switch, and the horizontal path in that row has a single color. Because there are no self-intersections, all strands crossing the row vertically have a color different from the horizontal path of the row. Removing the row splits the gadget into a part above and a part below the row, and the left and right endpoints of the horizontal path connect to different such parts. The strand containing the horizontal path of the row cannot be closed up without crossing the row, so the strand connects a terminal below and a terminal above the row. That is, it connects the bottom or left terminal to the top or right terminal. This contradicts our assumption that the bottom and left terminals lie on the same strand. ◀

We conclude:

► **Theorem 3.** *Given a curve arrangement and an integer k , it is NP-hard to decide whether there exists a sequence of k switches such that the resulting arrangement has no self-intersections.*



■ **Figure 12** Overlaying a grid. (a) None of the popular faces are the result of a self-intersection. (b) After superimposing a fine enough grid whose non-empty cells are of type (1) or (2), any popular face is now necessarily caused either by a self-intersection or a switch incident to two strands of the same colour.

Removing popular faces. Next, we consider the problem of removing all popular faces rather than only the self-intersections (that is, the aim is to produce a *basic* nonogram rather than an *advanced* one).

The idea is to globally overlay the construction from Section 4.1 with a sufficiently fine grid of perpendicular lines, in which none of the intersections with these new lines are switches. We construct our grid in such a way that each non-empty grid cell either has:

- (1) A single arc traversing it and connecting distinct sides of the cell.
- (2) Two arcs crossing exactly once (or meeting at a switch) and connecting opposite sides of the cell.

Additionally, we note that in our reduction of the self-intersection problem from the Permuter problem, configurations with no self-intersections have the added property that no switches involve two strands that belong to a single curve. To see this, it is enough to look at the design of 2-permuters: if both strands involved in their unique switch belonged to the same curve, then the upper crossing of their double-coil design would be a self-intersection. Because of this property, it is clear that once the grid is overlaid, all remaining popular faces are due to self-intersections (see Fig. 12). Therefore, in order to remove all popular faces we need to remove all self-intersections, and this is also sufficient.

We conclude:

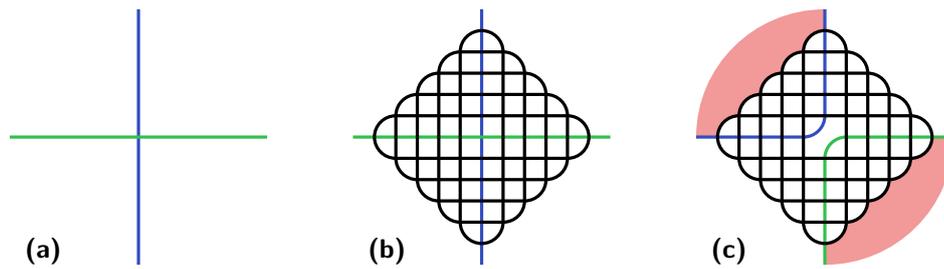
► **Theorem 4.** *Given a curve arrangement and a prescribed set of switches, it is NP-hard to decide whether it is possible to configure the switches such that the resulting arrangement has no popular faces.*

Removing popular faces with a minimum number of switches. Finally, we consider the setting where every intersection is an allowed switch. In this setting we wish to remove all popular faces using a minimum number of switches. The idea is similar to that in Section 4.1, but we will need to use a different gadget that ensures we cannot perform any switches (since having self-intersections does not necessarily imply there are no popular faces). To this end, we introduce the *Liège waffle*.

The Liège Waffle

In this section, we describe the *Liège waffle*. It is illustrated in Figure 13.

Essentially, we overlay c new closed curves on the two (crossing) terminal strands, such that each of the new curves is incident to each of the four unbounded faces, and that the disk bounded by the curve contains the crossing between the two terminal strands.



■ **Figure 13** The Liège waffle. (a) A crossing between two different curves. (b) We insert a sequence of c new curves. (c) Any sequence of local switches to emulate a global switch must be of length at least c or lead to a popular face.

► **Lemma 5.** *In a Liège waffle, any sequence of fewer than c switches must result in an arrangement with either the same combinatorial structure as the original, or at least one popular face.*

Proof. If we change the global connectivity, then one of the four unbounded faces will globally have two strands of the same curve; say the top left face (Figure 13 (c)). In order for this face to *not* be popular, these two strands must be consecutive along the curve. Initially, all intersections incident to the face are crossing, so they must all be uncrossed. There are $c + 1$ such intersections, so we need at least $c + 1$ switches. ◀

We conclude:

► **Theorem 6.** *Given a curve arrangement and an integer k , it is NP-hard to decide whether there exists a sequence of k switches such that the resulting arrangement has no popular faces.*

References

- 1 Yen-Chi Chen and Shun-Shii Lin. A fast nonogram solver that won the TAAI 2017 and ICGA 2018 tournaments. *ICGA Journal*, 41(1):2–14, 2019.
- 2 Phoebe de Nooijer. Resolving popular faces in curve arrangements. Master’s thesis, Utrecht University, 2022.
- 3 Phoebe de Nooijer, Soeren Nickel, Alexandra Weinberger, Zuzana Masárová, Tamara Mchedlidze, Maarten Löffler, and Günter Rote. Removing popular faces in curve arrangements by inserting one more curve. In *Proc. 38th European Workshop on Computational Geometry*, 2022.
- 4 Mees van de Kerkhof, Tim de Jong, Raphael Parment, Maarten Löffler, Amir Vaxman, and Marc J. van Kreveld. Design and automated generation of Japanese picture puzzles. *Comput. Graph. Forum*, 38(2):343–353, 2019.

4.2 Rotation Distance between Elimination Tree

Maïke Buchin ((Ruhr-Universität Bochum, Netherlands, maïke.buchin@rub.de)

Jean Cardinal (ULB – Brussels, Belgium, jean.cardinal@ulb.be)

Linda Kleist (TU Braunschweig, Germany, kleist@ibr.cs.tu-bs.de)

Boris Klemz (Universität Würzburg, Germany, boris.klemz@uni-wuerzburg.de)

Anna Lubiw (University of Waterloo, Canada, alubiw@uwaterloo.ca)

Torsten Mütze (University of Warwick – Coventry, UK, torsten.mutze@warwick.ac.uk)

Alexander Neuhaus (Ruhr-Universität Bochum, Germany, alexander.neuhaus-b7x@rub.de)

Lionel Pournin (Université Paris Nord – Villetaneuse, France, lionel.pournin@univ-paris13.fr)

License © Creative Commons BY 4.0 International license

© Maïke Buchin, Jean Cardinal, Linda Kleist, Boris Klemz, Anna Lubiw, Torsten Mütze, Alexander Neuhaus, Lionel Pournin

Problem description

Given a simple connected graph $G = (V, E)$, an *elimination tree* on G is obtained by selecting a root $r \in V$, and defining the subtrees of r as elimination trees on the connected components of $G \setminus \{r\}$. (Note that $G \setminus \{r\}$ can be connected, in which case r has a single child in the elimination tree.) There is a surjective mapping from the set of permutations S_n to elimination trees on an n -vertex graph: define r as the vertex of the current subgraph that has minimum index in the permutation [21]. Elimination trees are natural generalizations of binary search trees, which are obtained by letting G be a path on n vertices [38, 33].

Just like in binary search trees, one can define a *rotation* operation on elimination trees. Two elimination trees differ by a rotation if there exist two permutations that generate them, and which differ by a single adjacent transposition [50]. The rotation graph between elimination trees of a graph G is the skeleton of a polytope known as the *graph associahedron* of G [43, 44]. We consider the following computational problem:

Input: A simple connected graph G , two elimination trees T_1, T_2 on G , and an integer k .

Question: Can T_1 be transformed into T_2 using at most k rotations?

What is the complexity of this problem? Is it NP-hard?

The motivation here is the longstanding open question of the complexity of computing the rotation distance between two binary trees, a fixed-parameter version of the above problem in which G is an n -vertex path.

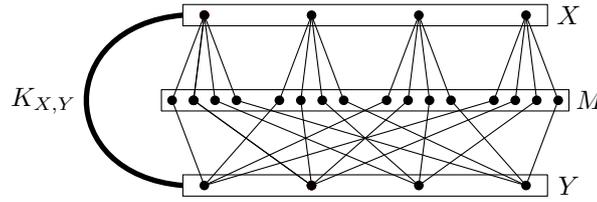
Directions

Complexity of tree-depth. The tree-depth $\text{td}(G)$ of a graph G is the minimum height of an elimination tree on G . Maybe the following result and its proof (see Pothen [56]) could be of some use in a hardness proof for the rotation distance.

► **Theorem 1.** *Deciding the tree-depth of a graph G is NP-hard.*

Proof. We reduce from the *complete balanced subgraph* problem in a bipartite graph. In this NP-hard problem, we wish to find the largest induced $K_{t,t}$ in a given bipartite graph G . Let us denote the maximum size t by $\text{CBS}(G)$. We claim that for the complement \bar{G} of a bipartite graph G ,

$$\text{td}(\bar{G}) = |V(G)| - \text{CBS}(G).$$



■ **Figure 14** Our gadget corresponding to an edge $\{x, y\}$.

Hardness follows from the claim. The claim can be proved by observing that elimination trees on the complement of a bipartite graph have a special structure: only one vertex has more than one children, and this vertex has exactly two children. Hence elimination trees have an “inverted Y” shape. The shortest branch has length $\text{CBS}(G)$. ◀

► **Conjecture 1.** *Rotation distance between elimination trees on a graph G is NP-hard, even if G is the complement of a bipartite graph.*

Edge gadget. In an attempt to prove NP-completeness, we considered the design of a gadget that allows to implement a binary choice.

Suppose we want to reduce from vertex cover. We need an edge gadget, such that the edge is covered when at least one of its endpoint is selected. The idea is to build a gadget graph G which has two sets of vertices X and Y and two elimination trees T_1 and T_2 for this graph, such that on any geodesic between T_1 and T_2 , all vertices in X are removed first (this is interpreted as “one endpoint of the edge is selected in the vertex cover”), or (this is a logical or, not an exclusive or) all vertices in Y are removed first (this is interpreted as “the other endpoint of the edge is selected in the vertex cover”). We need to design an edge gadget so that not removing X or Y first is very bad, and removing X or Y or $X + Y$ first is equally good and optimal, and any other order of removing them is not optimal.

Let us define the gadget \hat{G}_e for an edge e as follows, see also fig. 14: \hat{G}_e is a graph on two independent sets X and Y , each composed of m/k vertices, and a third independent set $M = \{1, 2, \dots, m\}$. In what follows, k is a function of m , to be decided later (maybe something like \sqrt{m}). The first vertex of X is connected to the first k vertices of M , the second to the next k vertices, and so on. The same goes for Y . Finally, we add all edges of the complete bipartite graph between X and Y , see fig. 14.

Now define two elimination trees T_1 and T_2 on this gadget. T_1 first eliminates all vertices of M in the order $1, 2, \dots, m$, then vertices of X . Then all vertices of Y are siblings in the tree. In T_2 , we first eliminate vertices of M in order $m, m - 1, \dots, 1$, then Y , then X .

► **Lemma 2.** *The rotation distance between T_1 and T_2 is $O(m^2/k)$.*

Proof. Move all vertices of X up. This costs (roughly)

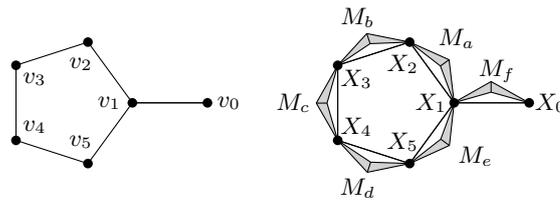
$$|X| \cdot |M| = \frac{n}{k} \cdot m$$

rotations. Now we are left with m/k connected components, each a star of $k + 1$ vertices. Now reorder the vertices of each star in $2k$ rotations each, costing

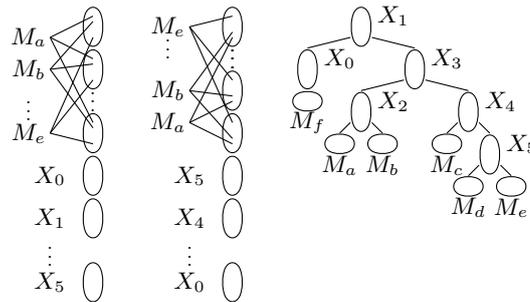
$$\frac{m}{k} \cdot 2k$$

rotations overall. Now push X back down, past all vertices of M , in again $\frac{m}{k} \cdot m$ rotations, taking care that the vertices of M are in reverse order (as in T_2). Finally, all vertices of X need to move down past all vertices of Y , costing $|X| \cdot |Y| = m^2/k^2$ rotations. The total is

$$2\frac{m}{k} \cdot m + \frac{m}{k} \cdot 2k + m^2/k^2 = 2\frac{m^2}{k} + o(m^2/k) \quad \blacktriangleleft$$



■ **Figure 15** Combining edge gadgets: graph.



■ **Figure 16** Combining edge gadgets: elimination tree.

It remains unclear whether copies of this gadget can be combined in a reduction; see Figures 16 and 16 for an example.

Reduction from token swapping. An alternative base problem would be the following (see Aichholzer et al. [55]). The input is a tree on n nodes, and two configurations (initial and final) where n distinct tokens are situated at the nodes. The goal is to get the tokens from their initial positions to their final positions (i.e. to effect the given permutation of tokens). The reconfiguration operation is a *swap* that exchanges the two tokens at the endpoints of an edge. In other words, we have the permutation group generated by the transpositions determined by the edges of the tree. The question is: Can you get from the initial to the final configuration with at most k swaps?

What are the similarities?

- working in a tree
- the reconfiguration graph is regular (from any configuration, there are $n - 1$ possible “flips”)
- the NP-hardness proof for token swapping on a tree uses ideas like the ones being suggested above: an “obvious” solution involves lots of flips and the only way to reduce the number of flips is to use a carefully structured solution that controls the movement.

Token swapping on a path is the same as rotation distance for elimination sequences of a clique. So they have an easy common case. The proof that token swapping on a tree is NP-complete is really long and hard [55].

Other open problems. The rotation distance problem can be specialized into possibly easier problems by restricting the types of possible input graphs G . The following cases might be relevant:

- G is a **star**. The rotation distance problem is in P (Lionel and Jean).
- G is a **path**. This is the famously open problem of computing the rotation distance between two binary search trees.
- G is a **broom**. Can this be proved polynomial-time equivalent to the case of a path?
- G is a **tree**. Same question as for brooms.

References

- 1 Wolk, E. The comparability graph of a tree. *Proceedings Of The American Mathematical Society*. **13** pp. 789-795 (1962)
- 2 Golumbic, M. Trivially perfect graphs. *Discrete Mathematics*. **24**, 105-107 (1978)
- 3 Yan, J., Chen, J. & Chang, G. Quasi-threshold Graphs. *Discrete Applied Mathematics*. **69**, 247-255 (1996)
- 4 Drange, P., Fomin, F., Pilipczuk, M. & Villanger, Y. Exploring the Subexponential Complexity of Completion Problems. *ACM Transactions On Computation Theory*. **7**, 14:1-14:38 (2015)
- 5 Nešetřil, J. & Mendez, P. On Low Tree-Depth Decompositions. *Graphs Comb.* **31**, 1941-1963 (2015)
- 6 Berendsohn, B. & Kozma, L. Splay trees on trees. *CoRR*. **abs/2010.00931** (2020)
- 7 Chapoton, F., Fomin, S. & Zelevinsky, A. Polytopal Realizations of Generalized Associahedra. *Canadian Mathematical Bulletin*. **45** pp. 537-566 (2002)
- 8 Del Pia, A. & Michini, C. On the diameter of lattice polytopes. *Discrete & Computational Geometry*. **55** pp. 681-687 (2016)
- 9 Deza, A., Manoussakis, G. & Onn, S. Primitive zonotopes. *Discrete & Computational Geometry*. **60** pp. 27-39 (2018)
- 10 Deza, A. & Pournin, L. Improved bounds on the diameter of lattice polytopes. *Acta Mathematica Hungarica*. **154** pp. 457-469 (2018)
- 11 Dotsenko, V., Shadrin, S. & Vallette, B. Toric varieties of Loday's associahedra and noncommutative cohomological field theories. *Journal Of Topology*. **12** pp. 463-535 (2019)
- 12 Kalai, G. & Kleitman, D. A quasi-polynomial bound for the diameter of graphs of polyhedra. *Bulletin Of The American Mathematical Society*. **26** pp. 315-316 (1992)
- 13 Klee, V. & Walkup, D. The d-step conjecture for polyhedra of dimension $d < 6$. *Acta Mathematica*. **117** pp. 53-78 (1967)
- 14 Mizera, S. Combinatorics and topology of Kawai-Lewellen-Tye relations. *Journal Of High Energy Physics*. pp. 2017:97 (2017)
- 15 Parlier, H. & Pournin, L. Flip-graph moduli spaces of filling surfaces. *Journal Of The European Mathematical Society*. **19** pp. 2697-2737 (2017)
- 16 Santos, F. A counterexample to the Hirsch conjecture. *Annals Of Mathematics*. **176** pp. 383-412 (2012)
- 17 Semple, C. & Steel, M. *Phylogenetics*. (Oxford University Press, 2003)
- 18 Sleator, D., Tarjan, R. & Thurston, W. Rotation distance, triangulations, and hyperbolic geometry. *Journal Of The American Mathematical Society*. **1** pp. 647-681 (1988)
- 19 Davis, M., Januszkiewicz, T. & Scott, R. Fundamental groups of blow-ups. *Advances In Mathematics*. **177**, 115-179 (2003)
- 20 Loday, J. Realization of the Stasheff polytope. *Archiv Der Mathematik*. **83**, 267-278 (2004,9)
- 21 Cardinal, J., Merino, A. & Mütze, T. Combinatorial generation via permutation languages. IV. Elimination trees. *CoRR*. **abs/2106.16204** (2021), <https://arxiv.org/abs/2106.16204>
- 22 Cardinal, J., Langerman, S. & Pérez-Lantero, P. On the Diameter of Tree Associahedra. *Electr. J. Comb.* **25**, P4.18 (2018)
- 23 Ceballos, C., Manneville, T., Pilaud, V. & Pournin, L. Diameters and geodesic properties of generalizations of the associahedron. *27th International Conference On Formal Power Series And Algebraic Combinatorics (FPSAC 2015)*. pp. 345-356 (2015,7)
- 24 Korte, B., Lovász, L. & Schrader, R. *Greedoids*. (Springer, Berlin, 1991)
- 25 Björner, A. & Ziegler, G. Introduction to greedoids. *Matroid Applications*. (1992)
- 26 Stasheff, J. Homotopy Associativity of H-Spaces. I. *Transactions Of The American Mathematical Society*. **108**, 275-292 (1963)

- 27 Tamari, D. Monoïdes préordonnés et chaînes de Malcev. (1951), Thèse de Mathématiques, Paris
- 28 Lee, C. The Associahedron and Triangulations of the n -gon. *European Journal Of Combinatorics*. **10**, 551-560 (1989)
- 29 Aspvall, B. & Heggernes, P. Finding minimum height elimination trees for interval graphs in polynomial time. *BIT Numerical Mathematics*. **34**, 484-509 (1994,12)
- 30 Deogun, J., Kloks, T., Kratsch, D. & Müller, H. On the vertex ranking problem for trapezoid, circular-arc and other graphs. *Discrete Applied Mathematics*. **98**, 39-63 (1999)
- 31 Ceballos, C., Santos, F. & Ziegler, G. Many non-equivalent realizations of the associahedron. *Combinatorica*. **35**, 513-551 (2015,10)
- 32 Pournin, L. The asymptotic diameter of cyclohedra. *Israel Journal Of Mathematics*. **219**, 609-635 (2017,4)
- 33 Pournin, L. The diameter of associahedra. *Advances In Mathematics*. **259** pp. 13-42 (2014)
- 34 Forcey, S., Lauve, A. & Sottile, F. New Hopf Structures on Binary Trees. *Discrete Mathematics & Theoretical Computer Science*. (2009,1)
- 35 Carr, M. & Devadoss, S. Coxeter complexes and graph-associahedra. *Topology And Its Applications*. **153**, 2155-2168 (2006)
- 36 Jordan, C. Sur les assemblages de lignes. *Journal Für Die Reine Und Angewandte Mathematik*. **70** pp. 185-190 (1869)
- 37 Demaine, E., Harmon, D., Iacono, J. & Pătraşcu, M. undefined.
- 38 Sleator, D. & Tarjan, R. Self-Adjusting Binary Search Trees. *J. ACM*. **32**, 652-686 (1985)
- 39 Nešetřil, J. & Mendez, P. Sparsity: Graphs, Structures, and Algorithms. (Springer,2012)
- 40 Katchalski, M., McCuaig, W. & Seager, S. Ordered colourings. *Discrete Mathematics*. **142**, 141 – 154 (1995)
- 41 Even, G. & Smorodinsky, S. Hitting sets online and unique-max coloring. *Discrete Applied Mathematics*. **178** pp. 71-82 (2014)
- 42 Postnikov, A., Reiner, V. & Williams, L. Faces of generalized permutohedra. *Documenta Mathematica*. **13** pp. 207-273 (2008)
- 43 Postnikov, A. Permutohedra, Associahedra, and Beyond. *International Mathematics Research Notices*. **2009**, 1026-1106 (2009)
- 44 Manneville, T. & Pilaud, V. Graph properties of graph associahedra. *Séminaire Lotharingien De Combinatoire*. **B73d** (2015)
- 45 Bixby, R., Cunningham, W. & Topkis, D. The Partial Order of a Polymatroid Extreme Point. *Mathematics Of Operations Research*. **10**, 367-378 (1985)
- 46 Diestel, R. Graph Theory. (Springer,2005)
- 47 Farber, M. & Jamison, R. Convexity in Graphs and Hypergraphs. *SIAM Journal On Algebraic Discrete Methods*. **7**, 433-444 (1986)
- 48 Bodlaender, H., Deogun, J., Jansen, K., Kloks, T., Kratsch, D., Müller, H. & Tuza, Z. Rankings of Graphs. *SIAM J. Discrete Math.* **11**, 168-181 (1998)
- 49 Bodlaender, H., Gilbert, J., Hafsteinsson, H. & Kloks, T. Approximating Treewidth, Pathwidth, Frontsize, and Shortest Elimination Tree. *J. Algorithms*. **18**, 238-255 (1995)
- 50 Bose, P., Cardinal, J., Iacono, J., Koumoutsos, G. & Langerman, S. Competitive Online Search Trees on Trees. *Proceedings Of The Annual ACM-SIAM Symposium On Discrete Algorithms (SODA20)*. (2020)
- 51 Schäffer, A. Optimal node ranking of trees in linear time. *Information Processing Letters*. **33**, 91-96 (1989)
- 52 Llewellyn, D., Tovey, C. & Trick, M. Local optimization on graphs. *Discrete Applied Mathematics*. **23**, 157-178 (1989)

- 53 Mozes, S., Onak, K. & Weimann, O. Finding an optimal tree searching strategy in linear time. *Proceedings Of The Nineteenth Annual ACM-SIAM Symposium On Discrete Algorithms, SODA 2008, San Francisco, California, USA, January 20-22, 2008*. pp. 1096-1105 (2008)
- 54 Onak, K. & Parys, P. Generalization of Binary Search: Searching in Trees and Forest-Like Partial Orders. *47th Annual IEEE Symposium On Foundations Of Computer Science (FOCS 2006), 21-24 October 2006, Berkeley, California, USA, Proceedings*. pp. 379-388 (2006)
- 55 Aichholzer, O., Demaine, E., Korman, M., Lynch, J., Lubiw, A., Masárová, Z., Rudoy, M., Williams, V. & Wein, N. Hardness of Token Swapping on Trees. *CoRR*. **abs/2103.06707** (2021), <https://arxiv.org/abs/2103.06707>
- 56 Pothen, A. The complexity of optimal elimination trees., Tech. Report CS-88-13, Pennsylvania State University
- 57 Karp, A. Bit Reversal on Uniprocessors. *SIAM Review*. **38**, 1-26 (1996)
- 58 Wilber, R. Lower Bounds for Accessing Binary Search Trees with Rotations. *SIAM J. Comput.* **18**, 56-67 (1989)
- 59 Demaine, E., Harmon, D., Iacono, J., Kane, D. & Patrascu, M. The geometry of binary search trees. *Proceedings Of The Twentieth Annual ACM-SIAM Symposium On Discrete Algorithms, SODA 2009, New York, NY, USA, January 4-6, 2009*. pp. 496-505 (2009)
- 60 Devadoss, S. A realization of graph associahedra. *Discrete Mathematics*. **309**, 271-276 (2009)
- 61 Dehornoy, P. On the rotation distance between binary trees. *Advances In Mathematics*. **223**, 1316-1355 (2010)
- 62 Emamjomeh-Zadeh, E., Kempe, D. & Singhal, V. Deterministic and probabilistic binary search in graphs. *Proceedings Of The 48th Annual ACM SIGACT Symposium On Theory Of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*. pp. 519-532 (2016)

4.3 Morphing Planar Graph Drawings Through 3D

Kevin Buchin (Technische Universität Dortmund, Germany, k.a.buchin@tue.nl)

Will Evans (University of British Columbia, Canada, will@cs.ubc.ca)

Fabrizio Frati (Roma Tre University, Italy, fabrizio.frati@uniroma3.it)

Irina Kostitsyna (TU Eindhoven, Netherlands, i.kostitsyna@tue.nl)

Maarten Löffler (Utrecht University, Netherlands, m.loffler@uu.nl)

Tim Ophelders (Utrecht University & TU Eindhoven, Netherlands, t.a.e.ophelders@uu.nl)

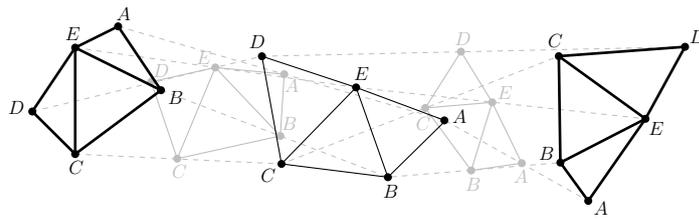
Carola Wenk (Tulane University, U.S.A., cwenk@tulane.edu)

Alexander Wolff (Universität Würzburg, Germany)

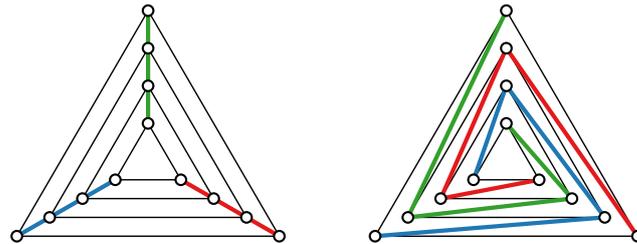
License © Creative Commons BY 4.0 International license

© Kevin Buchin, Will Evans, Fabrizio Frati, Irina Kostitsyna, Maarten Löffler, Tim Ophelders, Carola Wenk, Alexander Wolff

A *morph* is a continuous transformation between two given drawings of the same graph. A morph is required to preserve specific topological and geometric properties of the input drawings. For example, if the input drawings are planar and straight-line, the morph is required to preserve such properties throughout the transformation. A morphing problem often assumes that the input drawings are “topologically equivalent”, that is, they have the same “topological structure”. For example, if the input drawings are planar, they are required to have the same rotation system (i.e., the same clockwise order of the edges incident to each vertex) and the same walk bounding the outer face; this condition is obviously necessary (and it turns out, also sufficient) for a planarity-preserving morph to exist between the given drawings. A *linear* morph is a morph in which vertices move along straight-line segments, from their initial to their final position, at uniform speed. A *piecewise-linear* morph consists of a sequence of linear morphs, each of which is called a *step* (see Figure 17).



■ **Figure 17** A piece-wise linear morph consisting of two steps.



■ **Figure 18** Two planar straight-line drawings of the same n -vertex planar graph that require $\Omega(n)$ steps to be morphed into each other (in 2D).

We focused on the design of algorithms and bounds for morphing two-dimensional graph drawings in the three-dimensional space by means of few morphing steps. It is well-known that, given any two (topologically-equivalent) planar straight-line drawings of the same n -vertex planar graph, there exists a morph with $O(n)$ steps that transforms one drawing into the other one. In the plane, this bound is worst-case optimal [1].

Is it possible to reduce the number of steps by allowing the morph to exploit a third dimension? This question was first posed and studied by Arseneva et al. [3]. They proved that, given any two planar straight-line drawings of the same n -vertex tree, there exists a crossing-free morph with $O(\log n)$ steps that transforms one drawing into the other one. Whether a crossing-free morph with $o(n)$ steps exists for any two planar straight-line drawings of the same n -vertex planar graph is the main question that we addressed.

A Lower Bound

We soon realized that a major challenge is how to construct a three-dimensional morph with $o(n)$ steps between the two planar straight-line drawings that provide the lower bound for two-dimensional morphs, shown in Figure 18. In fact, we worked towards a proof of the following conjecture.

► **Conjecture 2.** *Every three-dimensional crossing-free morph between the planar straight-line drawings shown in Figure 18 requires $\Omega(n)$ steps, where n is the number of vertices of the graph.*

We came up with the following tentative approach for a proof of Conjecture 2. Consider two triangles T_1 and T_2 lying on horizontal planes in 3D, where T_1 is above T_2 . For $i = 1, 2$, let $a_i, b_i,$ and c_i be the vertices of T_i . Suppose that a_1 is connected to a_2, b_1 is connected to $b_2,$ and c_1 is connected to c_2 by means of three strings that spiral around each other $\Omega(n)$ times. This configuration can be reached with a single morphing step from the drawing in the right part of Figure 18: T_1 and T_2 are the outermost and innermost triangles, while the strings represent the colored paths. Then $\Omega(n)$ morphing steps seem to be necessary

to “despiralize” the colored paths, that is, to morph the described geometric object to a configuration in which the colored paths are vertical. This configuration can be reached with a single morphing step from the drawing in the left part of Figure 18. Similarly to Alamdari et al. [1], our strategy to prove the claimed lower bound consists of considering a concept of “winding number”, which is a measure of the described spiralization, and to prove that each morphing step changes the winding number only by a constant.

Conclusions

Although effective for tree drawings, the use of a third dimension does not seem to be helpful for morphing planar straight-line drawings with a sub-linear number of steps. On the other hand, we devised an approach that allows us to construct a morph between any two (possibly topologically non-equivalent) drawings of the same n -vertex planar graph with $O(n)$ steps. Such a morph does not exist when restricting to two dimensions. Immediate future work includes the formalization of this algorithm and the lower bound described above.

References

- 1 Soroush Alamdari, Patrizio Angelini, Fidel Barrera-Cruz, Timothy M. Chan, Giordano Da Lozzo, Giuseppe Di Battista, Fabrizio Frati, Penny Haxell, Anna Lubiw, Maurizio Patrignani, Vincenzo Roselli, Sahil Singla, Bryan T. Wilkinson, *How to Morph Planar Graph Drawings*, SIAM J. Comput. 46(2):824–852, 2017.
- 2 Patrizio Angelini, Michael A. Bekos, Fabrizio Montecchiani, Maximilian Pfister, *On Morphing 1-Planar Drawings*, 47th International Workshop on Graph-Theoretic Concepts in Computer Science (WG ’21): 270–282, 2021.
- 3 Elena Arseneva, Prosenjit Bose, Pilar Cano, Anthony D’Angelo, Vida Dujmović, Fabrizio Frati, Stefan Langerman, Alessandra Tappini, *Pole Dancing: 3D Morphs for Tree Drawings*, J. Graph Algorithms Appl. 23(3):579–602, 2019.
- 4 Giordano Da Lozzo, Giuseppe Di Battista, Fabrizio Frati, Maurizio Patrignani, Vincenzo Roselli, *Upward Planar Morphs*, Algorithmica 82(10): 2985–3017, 2020.

4.4 Statistics of Square Tiled Surfaces and the shearing block game

Vincent Delecroix (Université de Bordeaux, France, vincent.delecroix@u-bordeaux.fr)

Elise Goujard (Université de Bordeaux, France, elise.goujard@u-bordeaux.fr)

Luke Jeffreys (University of Bristol, UK, luke.jeffreys@bristol.ac.uk)

Hugo Parlier (University of Luxembourg, Luxembourg, hugo.parlier@uni.lu)

Saul Schleimer (University of Warwick, UK, s.schleimer@warwick.ac.uk)

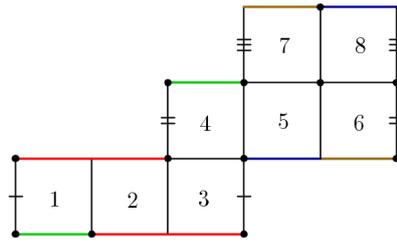
License  Creative Commons BY 4.0 International license

© Vincent Delecroix, Elise Goujard, Luke Jeffreys, Hugo Parlier, Saul Schleimer

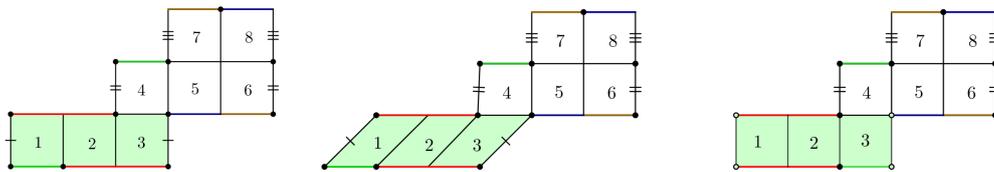
This working group focuses on some problems on square-tiled surfaces, as the shearing block game.

Discussed problems

► **Definition.** A *square-tiled surface* is an oriented compact connected surface obtained by gluing a finite number of isometric squares along parallel sides by translation (right \leftrightarrow left, up \leftrightarrow down).



■ **Figure 19** A square-tiled surface.



■ **Figure 20** A shear on the bottom left horizontal cylinder.

A cylinder on a square-tiled surface (more generally a translation surface) is a maximal collection of regular parallel closed geodesics. Cylinders are bounded by conical singularities of the induced flat metric on the square-tiled surface. On the picture above, the surface has 3 horizontal cylinders, and 2 vertical ones.

The “shearing block game” was suggested by Saul Schleimer and inspired by the talk of Hugo Parlier on playing puzzles on square-tiled surfaces.

The game goes as follows: Starting from one pattern, is it possible to get to another given pattern by a series of shears on cylinder blocks?

The layout problem

This game raises a first question: can every square-tiled surface be laid out in the plane?

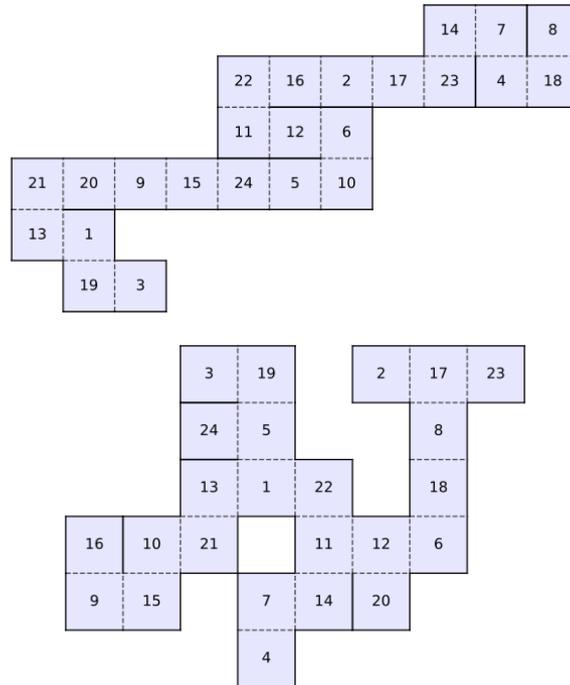
More precisely, a square-tiled surface made of n squares is conveniently encoded by a pair of permutations $(r, u) \in S_n \times S_n$ which record the gluing pattern: we glue to the right of the square labelled i the square $r(i)$ and above the square labelled i the square $u(i)$. It is also convenient to think of a square-tiled surface $(r, u) \in S_n \times S_n$ as an oriented 4-valent graph $G(r, u)$ where we have oriented edges for each pair $(i, r(i))$ and $(i, u(i))$ for $i \in \{1, 2, \dots, n\}$.

Given a square-tiled surface given by two permutations, how to draw it? We would ideally want to produce a layout in the plane such that

- i. the layout is made of unit squares centered on \mathbb{Z}^2 ;
- ii. all touching squares in the plane should be glued the same way in the surface; and
- iii. the layout is connected.

We will call such a layout a hard layout. If we relax condition ii. by cutting slits into the layout so that some adjacent squares are not adjacent in the surface then we will call such a layout a soft layout. Some possible variations on the layout are:

- soft layout (allowing slits) – It is equivalent to the existence of a spanning tree of $G(r, u)$ which is isomorphic to a subgraph of \mathbb{Z}^2 where the r edges are horizontal and the u edges are vertical (see the right of Figure 21);



■ **Figure 21** A snake layout (on the left, the extremities of the snake are the squares 3 and 8) , and a soft layout (on the right) of the surface

$$\begin{aligned}
 r &= (1, 22, 16, 2, 17, 23, 13)(3, 19)(4, 18)(5, 10, 21, 20, 9, 15, 24)(6, 11, 12)(7, 14) \\
 u &= (1, 5, 19)(2, 23, 14, 11, 22, 4, 7, 20, 12, 15, 10, 6, 18, 8, 17, 9, 16)(13, 21).
 \end{aligned}$$

- hard layout (do not allow slits) – It is equivalent to the existence of a connected subgraph of $G(r, u)$ which is isomorphic to an induced subgraph of \mathbb{Z}^2 where the r edges are horizontal and the u edges are vertical (see the right of Figure 23);
- snake layout – Impose that the spanning tree is a path (in particular, the graph $G(r, u)$ admits a non-oriented Hamiltonian path [4]) (see the left of Figure 21);
- staircase layout (only move east or north) – In terms of permutations we want an enumeration x_1, \dots, x_n of $\{1, 2, \dots, n\}$ so that two $x_{i+1} = r(x_i)$ or $x_{i+1} = u(x_i)$. It is equivalent to an oriented Hamiltonian path in the graph.

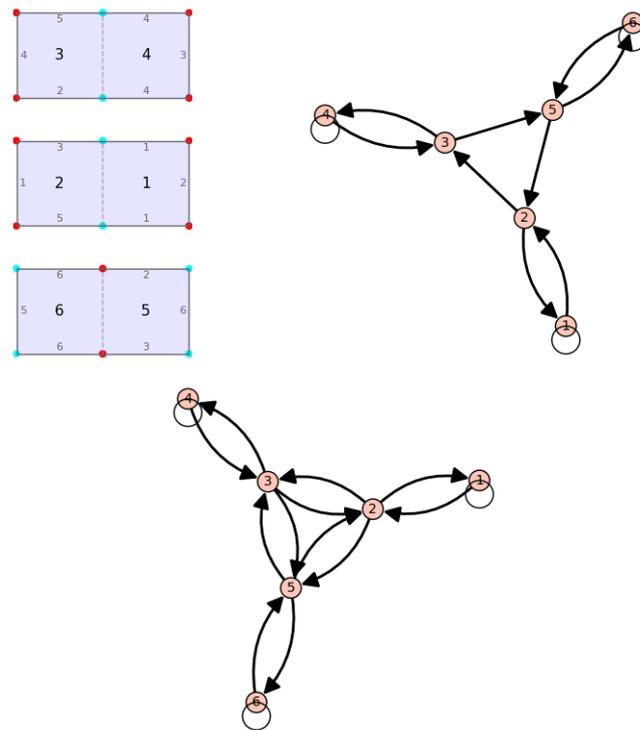
Here is an example lying in $\mathcal{H}(2, 2)^{odd}$ of a surface with no staircase nor snake layout (see Figure 22). The structure of the digraph associated to this square-tiled surface makes it clear that no staircase layout would be possible.

Some surfaces admit a snake layout but no staircase layout, see Figure 23.

There exist examples of square-tiled surfaces (r, u) with no soft layout. For instance the following square-tiled-surface made of 25 squares has no soft layout:

$$\begin{aligned}
 r &= (1, 6, 11, 16, 21)(2, 3, 4, 5)(7, 8, 9, 10)(12, 13, 14, 15)(17, 18, 19, 20)(22, 23, 24, 25) \\
 u &= (1, 2)(6, 7)(11, 12)(16, 17)(21, 22)
 \end{aligned}$$

It can be used as a “gadget” to be plugged in other surfaces but it is not a “generic” gadget as it imposes too many fixed points.



■ **Figure 22** The square-tiled surface and its associated digraphs (staircase and snake) for the permutations

$$\begin{aligned}
 r &= (1, 2)(3, 4)(5, 6) \\
 u &= (1)(2, 3, 5)(4)(6).
 \end{aligned}$$

Luke and Vincent coded up some methods that search for and, if one exists, then draws a staircase/snake/soft layout representation for a square-tiled surface. Luke’s method makes use of the `hamiltonian_path` method available inside Sage to find snake and staircase layouts, while Vincent’s method for finding soft and snake layouts frames the problem as an integer linear programming problem and utilises Sage’s `MixedIntegerLinearProgram` method. The code and some examples can be found in the files <https://coauthor.csail.mit.edu/file/ZwgDps62nDPygpfbW> and <https://raw.githubusercontent.com/videlec/Dagstuhl2022/master/Problem8-SquareTiledSurfaces/layout.py>.

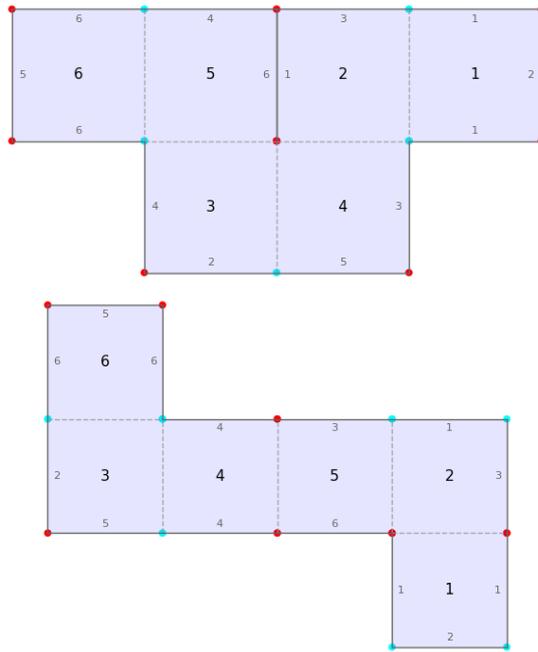
The structure that exists in the digraphs of the counter examples suggests a choice of “gadget” to include in higher complexity square-tiled surfaces.

► **Definition.** Let call a bridge a pair of edges e, e' such that

- e goes from u to v
- e' goes from v to u
- any path from u to v has to go through e
- any path from v to u has to go through e' .

► **Conjecture 1.** Consider the following quotient graph. Its vertices are the vertices adjacent to bridges in the initial graph and we put an edge between u and v if either

- there is a bridge between u and v



■ **Figure 23** Snake layouts for the surfaces

$$\begin{aligned}
 r &= (1, 2)(3, 4)(5, 6) & r &= (1)(2, 3, 4, 5)(6) \\
 u &= (1)(2, 3, 5, 4)(6) & u &= (1, 2)(3, 6, 5)(4).
 \end{aligned}$$

Both surfaces do not admit staircase layouts (although the layout on the right would be an east-south variation of a staircase layout).

■ *there is an oriented path between u and v that does not go through any bridge*
 Then a Hamiltonian path of the initial graph induces a Hamiltonian path on this quotient graph.

The general idea was that these bridge structures that appear in the counter-example can be used to build many more. However, one might expect that bridges are unlikely to be very common in general. In any case, we expect that having a Hamiltonian path in the digraph is not likely to be very common.

The experiments of this section lead to the following questions and conjectures:

Fix n and choose (uniformly at random) a pair of permutations r and u from the symmetric group S_n . You could also add conditions, say on the number of fixed points of the commutator $ru r^{-1}u^{-1}$ forcing the surface to be more flat/lower genus. This gives a “generic” square-tiled surface $S = S(r, u)$.

► **Question 1.** *As $n \rightarrow \infty$ does a random pair $(r, u) \in S_n \times S_n$ have a soft layout, a staircase layout, or (stronger still) a staircase layout starting from any of its squares?*

► **Conjecture 2.** *In that regime, the probability that a square-tiled surface has a soft/snake layout tends to 1, whereas the probability that a square-tiled surface has a positive staircase layout tends to 0.*

The intuition behind these conjectures is that generically r and u have very few short cycles, and these short cycles tend to be separated. This means that it is very hard to build a “gadget” (as the example with 25 squares). Here are some ideas for producing a (soft)

snake layout of generic S : We find all of the horizontal cylinders, sort them by length, layout the biggest cylinder, attach all of the medium and small cylinders to it (underneath) and then perform some kind of snake layout with the remaining long cylinders. If this works, it may help promoting from the soft layout to the hard.

Also, generically the big cycle of say r is too big, so too many cycles (if u) should meet it, making it impossible to get a positive staircase layout.

► **Conjecture 3.** *Deciding whether a given pair $(r, u) \in S_n \times S_n$ has a soft/hard/staircase-/snake layout is NP-complete.*

► **Question 2.** *For a random square-tiled surface and a random spanning tree of that square-tiled surface, what is the area of the obtained (possibly overlapping) layout? What is the diameter of the layout? What is the diameter of the spanning tree? What proportion of the area corresponds to the overlapping part of the layout?*

The shearing block game

The first question raised for the shearing block game was the classification of the orbits for the shearing block moves. Obviously it does not change the total number of squares, nor the stratum (characterized by the order of conical singularities of the flat metric) the surface belongs to. One can check that it also doesn't change the connected component of this stratum.

► **Conjecture 4.** *The following invariants classify orbits of the shearing block game:*

- *connected component of stratum (Kontsevich-Zorich)*
- *number of squares n*

This conjecture was checked by computer experiments (see code at <https://github.com/videlec/Dagstuh12022>) in the following cases:

- genus 2:
 - $\mathcal{H}(2)^{hyp}$ up to 30 squares
 - $\mathcal{H}(1, 1)^{hyp}$ up to 18 squares
- genus 3:
 - $\mathcal{H}(4)^{hyp}$ up to 12 squares
 - $\mathcal{H}(4)^{odd}$ up to 10 squares
 - $\mathcal{H}(3, 1)^c$ up to 10 squares
 - $\mathcal{H}(2, 2)^{hyp}$ up to 10 squares
 - $\mathcal{H}(2, 2)^{odd}$ up to 10 squares
 - $\mathcal{H}(2, 1, 1)^c$ up to 10 squares
 - $\mathcal{H}(1, 1, 1, 1)^c$ up to 10 squares

If the conjecture is correct, the size of each graph (connecting the different configurations of a game) is computable in the following way. For each stratum component \mathcal{C} there exists a quasi-modular form whose n th coefficient is the number of square-tiled surfaces in \mathcal{C} made of n squares.

Experimental data for the diameter could also be obtained from surface-dynamics, but first, let us define formally the graph associated to the game. An edge could represent several possible moves:

- 1) one shear in one cylinder ($1/\ell$ fractional Dehn twist where ℓ is the perimeter of the cylinder)
- 2) any number of shears in one cylinder
- 3) any number of shears in any number of (parallel) cylinders
- 4) 90 degree rotation

Numerical experiments give the following sizes and diameters for the graphs with respect to the choices 1) or 4) for the edges:

n	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
size	3	9	27	45	90	135	201	297	405	525	693	918	1062	1395	1620	2043	2295	2835	3120	3915	4158	5085	5337
diam.	2	4	13	12	15	13	18	17	19	18	19	19	21	20	21	21	22	21	23	22	24	24	23

Furthermore, we can consider various metrics on these graphs. The initial metric considered was the natural metric for “slow moves”: an edge representing a single shear of size one (right or left) in a single cylinder as length 1. There are various suggestions of other costs:

- $\log(k + 1)$ for a shear of size k in a single cylinder,
- $\log(k + 1)$ for a shear of size k on a “stack” horizontal cylinders,
- $\log(k + 1)$ for a k fold shear in any collection of horizontal cylinders
- 1 for any horizontal shear in any collection of horizontal cylinders. . .

As we change the metric, the diameter of the graph changes. This raises the following question:

► **Question 3.** *Is there a (natural) definition of a “single move” and a corresponding metric on the graph that allows to connect two vertices very quickly? Is it possible that two “large” moves suffice?*

Some leads considered to solve the conjecture were the following:

- one-cylinder surfaces are easy to play with (in one direction), and we know from [3] how many such surfaces we have in a given stratum.
- the $\mathrm{SL}(2, \mathbb{Z})$ orbit closures are connected via the shearing block moves.

This raises the following question:

► **Question 4.** *How do $\mathrm{SL}(2, \mathbb{Z})$ orbit closures lie inside of the graph? How do they “approach” each other?*

Here are further questions concerning this game played on random surfaces:

- **Question 5.** *Consider a random square-tiled surface S in the same setting as before.*
 - *What is the size of the component (of the shearing block game) containing a generic S ? (Conjecture: $(N!)^2 / N^{\log(n)}$)*
 - *What is the diameter of the component? Can we navigate?*
 - *Suppose that we perform random moves to S to get the sequence of surfaces (S_k) with associated permutations r_k and u_k . Note that the cycle type of the commutator $r_k u_k R_k U_k$ is fixed. However, we can still ask: does r_k converge to the generic permutation as k tends to infinity?*
 - *How quickly does the random walk (on the component) mix?*

An application – if the graph is connected, is this a good way to sample square-tiled surfaces?

References

- 1 Schloss Dagstuhl – Editorial Office, *The dagrep-v2018 class*. Schloss Dagstuhl, Germany, 2011.
- 2 John Q. Open; Joan R. Access, *Seminar Sample*, Dagstuhl Reports, 1:1, 1–8, 2011.
- 3 V. Delecroix, E. Goujard, P. Zograf, A. Zorich, *Contribution of one-cylinder square-tiled surfaces to Masur-Veech volumes. With an appendix by Philip Engel*. Astérisque No. 415, *Quelques aspects de la théorie des systèmes dynamiques: un hommage à Jean-Christophe Yoccoz. I* (2020), 223–274. <https://arxiv.org/abs/1903.10904>
- 4 <https://www.cs.utoronto.ca/~gbrunet/project2.pdf>

4.5 The graph of minimal configurations of a multicurve

Mark Bell (*Hampshire, UK*)

Benjamin Burton (*The University of Queensland – Brisbane, Australia, bab@maths.uq.edu.au*)

Erin Moriarty Wolf Chambers (*St. Louis University, US, erin.chambers@slu.edu*)

Vincent Delecroix (*University of Bordeaux, France, vincent.delecroix@labri.fr*)

Niloufar Fuladi (*University Paris-Est – Marne-la-Vallée, France, niloufar.fuladi@aol.com*)

Luke Jeffreys (*University of Bristol, UK, luke.jeffreys@bristol.ac.uk*)

Francis Lazarus (*CNRS – Grenoble, France, Francis.Lazarus@grenoble-inp.fr*)

Arnaud de Mesmay (*University Paris-Est – Marne-la-Vallée, France, ademesmay@gmail.com*)

Hugo Parlier (*University of Luxembourg, Luxembourg, hugo.parlier@uni.lu*)

Saul Schleimer (*University of Warwick – Coventry, UK, s.schleimer@warwick.ac.uk*)

Éric Colin de Verdière (*CNRS & LIGM – Marne-la-Vallée, France, ericcolindeverdiere@u-pem.fr*)

Hsien-Chih Chang (*Dartmouth College – Hanover, US, hsien-chih.chang@dartmouth.edu*)

Jean Cardinal (*ULB – Brussels, Belgium, jean.cardinal@ulb.be*)

License © Creative Commons BY 4.0 International license

© Francis Lazarus, Mark Bell, Benjamin Burton, Erin Moriarty Wolf Chambers, Vincent Delecroix, Niloufar Fuladi, Luke Jeffreys, Francis Lazarus, Arnaud de Mesmay, Hugo Parlier, Saul Schleimer, Éric Colin de Verdière, Hsien-Chih Chang, Jean Cardinal

This working group focused on the graph of minimal configurations of a multicurve.

Discussed Problem

Consider a set of curves (a.k.a a *multicurve*) Γ on a surface S . Here, we consider curves up to continuous deformations (free homotopy), so that a curve is actually a conjugacy class in $\pi_1(S)$. We shall only consider **primitive** curves, that is curves that are not proper powers of other curves, moreover pairwise distinct in Γ . A **configuration** of Γ is a choice of a representative of each curve in Γ so that all their intersections on S are transverse and there is no triple points. A configuration is considered up to isotopy. A configuration is **minimal** if its number of double points is minimal among all possible configurations of Γ .

It is known (see e.g. [10]) that any configuration of Γ can be brought to a minimal one by elementary moves, a.k.a. shadows of Reidemeister moves. In fact, it was shown that any two configurations of Γ are related by a *monotonic* sequence of moves, where the number of double points changes monotonically [3, 13, 2]. In particular, any two minimal configurations of Γ are related by a sequence of 3-3 moves. The **configuration graph** of Γ has for vertex set the set of minimal configurations of Γ and two vertices are connected by an edge if they are related by a single 3-3 move. The previous remark ensures that this graph is connected. It follows from Hass and Scott [11] that this graph is finite. There are two natural questions related to this graph:

- What is its size?
- What is its diameter?

Following Souto and Vo [14], we can reduce to the case where Γ is *filling*, i.e. such that each of its configurations cuts S into topological disks and annuli bounded on one side by a boundary component of S . Indeed, after putting Γ into minimal configuration [2], one can construct a subsurface $S(\Gamma) \subset S$ where Γ is filling. $S(\Gamma)$ is obtained by replacing every component C of the complement $S \setminus \Gamma$ that is not a disk by a set of annuli, with one annulus per boundary component of C that is not a boundary component of S . (Formally, one needs to replace C by its metric completion in order to speak of its boundary components). We call

such an annulus a **connecting annulus**. The fact that all minimal configurations are related by 3-3 moves implies that the topology of $S(\Gamma)$ is independent of the minimal configuration used to cut S . Now, two multicurves Γ and Γ' have the same type in S (see below for a definition) if and only if there is a homeomorphism $\varphi : S(\Gamma) \rightarrow S(\Gamma')$ sending Γ to Γ' (up to homotopy) such that φ can be extended to a self-homeomorphism of S . This last condition can be checked easily as follows: Consider the graph G_Γ whose vertices are the connecting annuli in $S(\Gamma)$ and whose edges correspond to connecting annuli bounding a same component of $S \setminus S(\Gamma)$. We moreover mark every vertex of G_Γ with the topological type of the incident component in $S \setminus S(\Gamma)$ (its genus and its total number of boundary components). Then φ extends to a self-homeomorphism of S if and only if it induces an isomorphism between the marked graphs G_Γ and $G_{\Gamma'}$.

We can thus assume that Γ is filling. In this case any configuration determines a combinatorial map in the obvious way: the vertex and edges of this map are simply given by the arrangement of the curves in the configuration, and the faces of the map correspond to the complementary disks and annuli (one per boundary component of S). These maps are 4-regular as we assume that there is no triple point in a configuration. We can replace the minimal configurations in the configuration graph by their associated combinatorial maps. Note that applying a self-homeomorphism of S to a configuration does not change (the isomorphism class of) its associated combinatorial map. Also note that the configuration graph is entirely determined by *any* of its configuration maps, from which we can recover all the other configurations by 3-3 moves. Say that two multicurves have the same **type** if they are in the same orbit of the mapping class group. More formally, this means that there is a 1-1 correspondence between the two sets of free homotopy classes of curves in the two multicurves that is induced by some self-homeomorphism of S . Hence, two filling multicurves have the same type if and only if they have the same configuration graph.

Recently, Souto and Vo [14] described a polynomial time algorithm to detect when two curves have the same type. Using normal coordinates with respect to a fixed triangulation, their algorithm enumerates a polynomial number of candidate mapping classes that must contain a mapping class sending one curve to the other one in case they indeed have the same type. It remains to check whether any candidate mapping class relates the input curves to decide if they have the same type. Our study of the graph of minimal configurations is motivated by an alternative approach to the algorithm of Souto and Vo.

Looking for hyperbolic configurations

Suppose that our surface S is equipped with some Riemannian metric. Freedman et al. [8] proved that shortening a given curve as much as possible for this metric puts the curve in minimal configuration¹. Conversely, Neumann-Coto [12] proved that every minimal configuration of a multicurve is in the configuration of shortest geodesics for *some* Riemannian metric. It thus seems natural to encode a configuration by a metric for which it is a minimizer. Hyperbolic metrics provide an interesting subset of metrics. We say that a configuration of a multicurve is **hyperbolic stretchable** or, more simply, **hyperbolic**, if it can be realized by homotopic geodesics for some hyperbolic metric. Hass and Scott [11] gave counterexamples to the fact that every multicurve configuration is hyperbolic. It follows that the graph of hyperbolic configurations is in general smaller than the whole graph of configurations. The consideration of hyperbolic configurations raises several questions.

¹ In full generality, one should allow the minimal configurations to have crossings with multiplicity larger than two. A crossing with multiplicity k should count for $\binom{k}{2}$ simple crossings.

1. One already serves as a conclusion in [4]: Given a multicurve, is there an algorithm to (construct or) detect configurations that are hyperbolic?
2. Is the graph of hyperbolic configurations of a multicurve connected?
3. How big is the graph of hyperbolic configurations compared to the whole graph of configurations? What is its diameter?

The answer to the second question should be positive: simply interpolate hyperbolic metrics associated to two configurations in the Teichmüller space of S . The interpolation path should be generic in the sense that singular metrics (for which the multicurve has crossings of multiplicity larger than two) along this path should be isolated and should allow for only one non-regular crossing of multiplicity exactly three (this requires an argument to claim that this is indeed the generic situation.). Then, the regular configurations along this path are related by 3-3 moves, showing the connectivity of the graph of hyperbolic configurations.

We have not studied the third question very much. What should be clear is that the whole configuration graph can be very big with respect to the number of crossings. As an indication one may consider the number $\Omega(2^{n^2/5})$ of pseudoline arrangements [5], where the number of crossings is quadratic with respect to the number n of pseudolines.

We now turn to the first question. Unfortunately, determining whether a configuration is hyperbolic stretchable seems quite hard. We actually prove.

► **Proposition 2.** *Given a combinatorial map representing a configuration of a multicurve on a surface, it is $\exists\mathbb{R}$ -hard to decide if the configuration is hyperbolic.*

Recall that a problem is $\exists\mathbb{R}$ -hard if the Existential Theory of the Reals (ETR) many-one reduces to it in polynomial-time, and that the problems in ETR are emptiness of semi-algebraic systems.

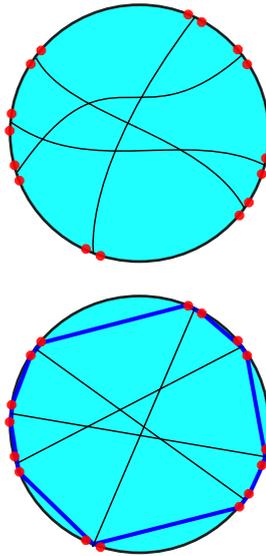
Proof. The proof goes by showing that the stretchability of a pseudoline arrangement in the plane reduces in polynomial time to the hyperbolic stretchability of a multicurve configuration. The proposition then follows from Mnëv’s universality theorem implying that the problem of stretchability of pseudoline arrangements is polynomially equivalent to ETR. See [7] or a quick overview in [6] for precise statements. Recall that a pseudoline arrangement is a collection of simple arcs in a disk such that the arc endpoints are on the disk boundary and such that any two arcs intersect exactly once and transversely.

We give two reductions, transforming an arrangement of n pseudolines into a configuration of n curves either on a sphere with $4n$ punctures or on a closed surface of genus $2n$.

Reduction to a configuration on a punctured sphere. Consider an instance I_1 of stretchability of an arrangement of n pseudolines. We build an instance I_2 of hyperbolic stretchability of a multicurve configuration such that I_1 is positive iff I_2 is.

Let D be the disk containing the instance I_1 . We introduce $4n$ punctures on the boundary of D in pairs; each pair surrounds one of the $2n$ endpoints of the n pseudolines as on the next figure. We then take a copy D' of D and attach them together: Think of D as the Northern hemisphere, D' as the Southern hemisphere, the boundary of the disks to be the equator. The punctures coming from the disks are identified so that there are $4n$ punctures in total. Each pseudoline is now a closed curve, passing once in the Northern and once in the Southern hemispheres. The reduction takes trivially polynomial time.

Assume that I_1 is positive. We have an arrangement of lines in the Euclidean disk. We put the $4n$ punctures on the boundary of the disk, and then consider the $4n$ -gon whose vertices are the punctures (so we cut off some small pieces of the disk). Now we view this



$4n$ -gon as an ideal (hyperbolic) polygon induced by the Klein model of the disk. We take a copy of this ideal polygon and attach them side by side. This gives a hyperbolic metric on the $4n$ -punctured sphere where the arrangement of the closed curves in I_2 is realized. (This is because, when we cut off the disk into a $4n$ -gon, no intersection between pseudolines has been removed.) We then relax these closed curves to geodesics in their respective homotopy classes. If the two punctures surrounding an endpoint of a pseudoline were close enough together, this doesn't change the combinatorial arrangement of the curves, because the initial pseudoline arrangement was in general position. So I_2 is positive. (We know that we need doubly exponential precision for the location of the punctures [9], but we don't care: after all we just consider decision problems!)

Conversely, assume that I_2 is positive. Consider the corresponding hyperbolic surface with cusps, S . In S , consider a topological disk D that contains I_1 as a topological arrangement. Consider the Klein model K of the universal cover of S . Lift D (and the pieces of the lines inside D) in K . The picture in K is a topological disk \tilde{D} and line segments \tilde{D} whose combinatorial arrangement is exactly the input to I_1 . Extend these lines to straight line segments with endpoints on the boundary of \tilde{D} ; because, in \tilde{D} , every pair of segments already cross (since I_1 is a pseudoline arrangement), this extension does not create new crossings. In other words, if we now view \tilde{D} as a Euclidean disk, our arrangement is now combinatorially equivalent to I_1 . So I_1 is positive.

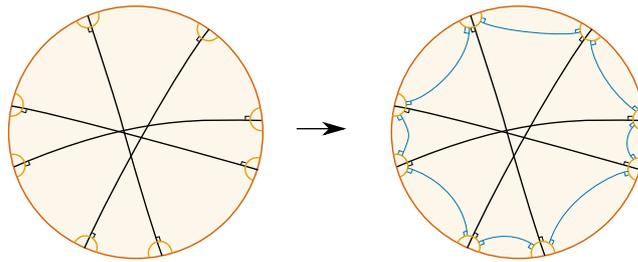
Reduction to a configuration on a closed surface. As before we start with an instance I_1 of an arrangement of n pseudolines ℓ_1, \dots, ℓ_n in a disk D . As before, we introduce $4n$ punctures on the boundary ∂D of D surrounding the $2n$ pseudoline endpoints. This divides ∂D into $4n$ arcs, where one out of every two contains an endpoint. We denote by α_i the arcs that contains an endpoint and by β_i the remaining ones. We then take a copy D' of D and glue their arcs α_i via the identity map. This results in a sphere with $2n$ boundaries. We cap off each of these boundaries with a one-holed torus to obtain a configuration of n closed curves on a closed surface of genus $2n$. This ends the construction of our instance I_2 for

hyperbolic stretchability². The construction trivially takes polynomial time.

Assume that I_1 is positive. We have an arrangement of L lines in the Euclidean disk. Considering the Klein model of the disk this provides an isomorphic arrangement of hyperbolic lines that we still denote by ℓ_1, \dots, ℓ_n . Let α_i be a collection of geodesics as follows.

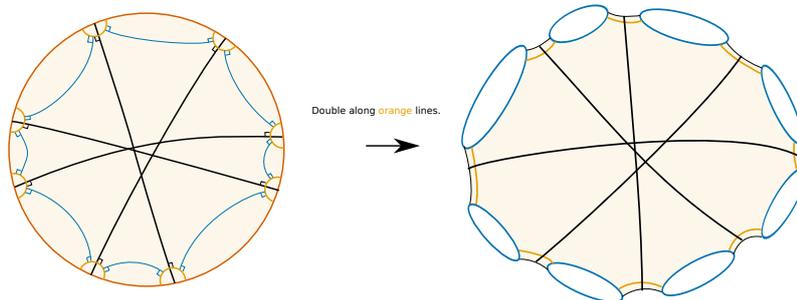
1. If $i \neq j$ then α_i is disjoint from α_j (even at infinity – they do not share an ideal point).
2. α_i and ℓ_i meet in a single point.
3. α_i cuts a half plane H_i off of D ; the intersection of H_i with the line arrangement L is exactly one end of ℓ_i .

Let β_i be the common perpendicular to α_i and α_{i+1} .



Claim: β_i is disjoint from L .

We now cut D along the α_i and β_i to obtain a hyperbolic right-angled $4n$ -gon. We double this $4n$ -gon and glue the two copies along the α_i . We obtain a hyperbolic sphere with $2n$ geodesic boundaries as shown below.

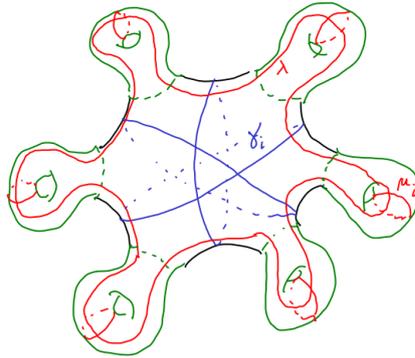


We finally cap off the $2n$ geodesic boundaries with hyperbolic one holed tori. We denote by S the resulting hyperbolic surface. Let γ_i be the union of ℓ_i and its copy; this is a closed geodesic on S . Clearly, S and the γ_i certify that I_2 is positive.

The reverse implication goes exactly as in the previous reduction on a punctured sphere. This ends the proof of the proposition.

We can slightly strengthen the proposition by restricting the hardness result to filling configurations. In order to do so we just need to add $2n + 1$ curves to the γ_i in the previous construction. One of these curves, λ , goes through all the handles as on the next figure. And the $2n$ other curves, μ_i , cut open each handle. The μ_i are disjoint from the γ_i and will thus be so in any minimal configuration. This allows us to recover the extra $2n + 1$ curves without the need for a marking. ◀

² An alternative construction consists of taking four copies of D , gluing (cyclically) the j th and $(j + 1)$ th copies along the α_i or along the β_i according to the parity of j . This gives a surface of genus $2n - 1$ with a configuration of $2n$ (?) curves.



In search of a canonical configuration

In view of the hardness result in Proposition 2, the restriction of the configuration graph to hyperbolic configurations might not be so beneficial from the computational viewpoint. There is however a surprising benefit to consider hyperbolic configurations. They contain a preferred, i.e. canonical, configuration (at the expense of allowing degenerate configurations with crossings of multiplicity higher than two)!

Suppose that S is a (closed, connected, oriented) surface. Suppose that γ is a (simple?) closed curve in S . For any hyperbolic metric σ on S , we can define $l_\gamma(\sigma)$ to be the length of the geodesic representative of γ with respect to σ . Thus l_γ is a function from Teichmüller space to \mathbb{R} . Wolpert proves that l_γ is convex with respect to the Weil-Petersson metric [15] and Bestvina et al. with respect to some well chosen Fenchel-Nielsen coordinates [1].

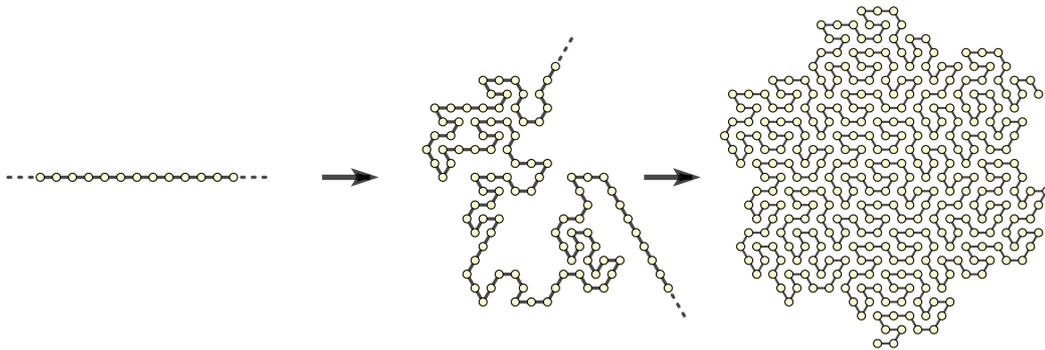
Suppose now that (γ_i) is a filling collection of curves. Then there is a *unique* σ that minimises the sum $\sum_i l_{\gamma_i}$. The intersection pattern of the geodesic representatives of the γ_i , with respect to σ can therefore be considered a canonical form for the family γ_i . Note however that there might be points with more than double intersections in this canonical form.

We conclude with many questions on this minimising metric.

1. Is the minimising metric algebraic? How do the degree/height of the algebraic numbers depend on the complexity of the given curves (γ_i) ?
2. Given a hyperbolic metric σ , can we find a collection of curves (γ_i) having σ as a minimiser? Written this way the answer is no (countable versus uncountable). We hence have two questions: – Is any hyperbolic metric σ the minimizer of some *weighted* collection of curves? Of simple curves? Of two simple curves? – What are the hyperbolic metrics σ which are minimizers of a collection of curves?
3. Is there such a minimiser in the flat setting? Note that the length function is convex in the $SL(2, \mathbb{R})$ directions. However, it is hard to believe that it is convex along any linear deformation (in period coordinates).
4. What happens in the case of the torus (hyperbolic or flat)?
5. Is there a way to get a nice bound on the diameter of the configuration graph using this approach? A discrete algorithm to find the moves?
6. Perhaps Whitehead's algorithm is relevant here? We can use Whitehead to find, in polynomial time, an optimal cut system (cutting the surface into a connected planar surface). See the exposition of Berge; he describes how to use max-flow-min-cut to search for the Whitehead automorphisms.
7. There is a combinatorial proof of Nielsen realisation, I believe due to Hensel, Osajda, and Przytycki.

References

- 1 Mladen Bestvina, Ken Bromberg, Koji Fujiwara, and Juan Souto. Shearing coordinates and convexity of length functions on teichmüller space. *American journal of Mathematics*, pages 1449–1476, 2013.
- 2 Hsien-Chih Chang and Anaud de Mesmay. Tightening curves on surfaces monotonically with applications. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 747–766. SIAM, 2020.
- 3 Maurits de Graaf and Alexander Schrijver. Making curves minimally crossing by Reidemeister moves. *Journal of Combinatorial Theory, Series B*, 70(1):134–156, 1997.
- 4 Vincent Despré and Francis Lazarus. Computing the geometric intersection number of curves. *Journal of the ACM (JACM)*, 66(6):1–49, 2019.
- 5 Adrian Dumitrescu and Ritankar Mandal. New lower bounds for the number of pseudoline arrangements. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 410–425. SIAM, 2019.
- 6 Jeff Erickson. Optimal curve straightening is $\exists\mathbb{R}$ -complete.
- 7 Stefan Felsner and Jacob E Goodman. Pseudoline arrangements. *Handbook of Discrete and Computational Geometry*, 3, 2017.
- 8 Michael Freedman, Joel Hass, and Peter Scott. Closed geodesics on surfaces. *Bull. London Math. Soc.*, 14:385–391, 1982.
- 9 Jacob E Goodman, Richard Pollack, and Bernd Sturmfels. Coordinate representation of order types requires exponential storage. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 405–410, 1989.
- 10 Joel Hass and Peter Scott. Intersections of curves on surfaces. *Israel Journal of Mathematics*, 51(1-2):90–120, 1985.
- 11 Joel Hass and Peter Scott. Configurations of curves and geodesics on surfaces. *Geometry and Topology Monographs*, 2:201–213, 1999.
- 12 Max Neumann-Coto. A characterization of shortest geodesics on surfaces. *Algebraic and Geometric Topology*, 1:349–368, 2001.
- 13 J.M. Paterson. A combinatorial algorithm for immersed loops in surfaces. *Topology and its Applications*, 123(2):205–234, 2002.
- 14 Juan Souto and Thi Hanh Vo. Deciding when two curves are of the same type.
- 15 Scott A. Wolpert. Convexity of geodesic-length functions: a reprise.



■ **Figure 24** A turning machine folding into a Gosper curve.

4.6 Turning Machines

Kevin Buchin (Technische Universität Dortmund, Germany, k.a.buchin@tue.nl)

Jean Cardinal (Université Libre de Bruxelles, Belgium, jean.cardinal@ulb.be)

Will Evans (University of British Columbia, Canada, will@cs.ubc.ca)

Irina Kostitsyna (TU Eindhoven, the Netherlands, i.kostitsyna@tue.nl)

Alexander Wolff (Universität Würzburg, Germany)

License © Creative Commons BY 4.0 International license

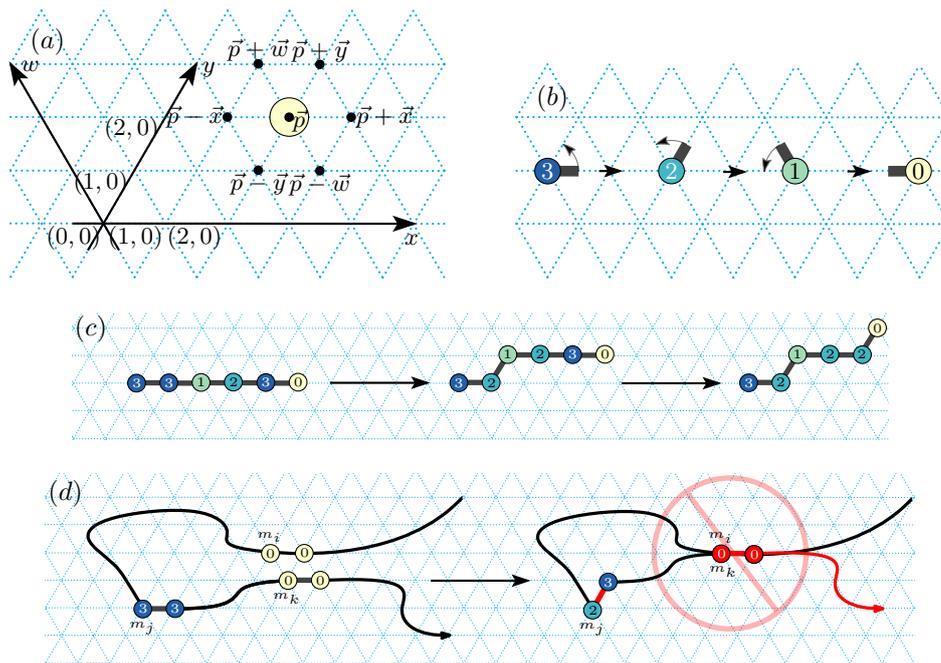
© Kevin Buchin, Jean Cardinal, Will Evans, Irina Kostitsyna, Alexander Wolff

The *turning machine* [1] is a very simple model of a molecular robot whose task is to fold into a desired shape. Such a machine consists of a number of molecular computing units (called *monomers*) that together form a chain. Depending on their states, the monomers can rotate, thus bending the chain and changing its global structure (see Figure 24). On his website [2], Woods gives an informal description of the model and has example videos where turning machines fold into certain shapes. During the workshop, we discussed how to assign initial states to the monomers such that they can then rotate and bring the chain into a prescribed shape.

Model. Consider a triangular grid $G_\Delta = (V_\Delta, E_\Delta)$. A turning machine T is a chain of monomers in G_Δ . For simplicity, we introduce a system of coordinates on G_Δ (see Figure 25). Initially, the monomers of a turning machine form a horizontal chain m_1, m_2, \dots, m_n that starts from the origin and extends East, that is, monomer m_i occupies grid point $(i - 1, 0)$. We orient the edges of the chain from left to right. For $i \in \{1, 2, \dots, n - 1\}$, monomer m_i has state s_i , which is an integer value that indicates how many times the edge leaving m_i needs to rotate.

A turning machine folds by the means of local rotations of monomers. In one step, a monomer m_i with a non-zero state s_i can rotate its outgoing edge by 60° – counterclockwise if $s_i > 0$ and clockwise if $s_i < 0$. This results in the suffix of the chain translating by a unit vector in the triangular grid. We require the chain to not self-intersect, so not all monomers can rotate at any moment in time. In other words, some monomers can be *blocked*. When a monomer rotates, the absolute value of its state decreases by one.

For example, the second monomer from the left in Figure 25(c) can rotate. As a result, the state of this monomer decreases from 3 to 2 and the subchain to the right of the monomer translates by the unit vector shown in the figure.



■ **Figure 25** Figure taken from [1]. Turning machine model. (a) Triangular grid conventions. A configuration showing a single monomer on the triangular grid G_Δ , along with axes x , y and w . (b) Local movement (rotation): A monomer in state 3 pointing to the east undergoes three turning rule applications finishing in state 0 with no more applicable rules. Locally, the monomer exhibits a rotation motion. (c) Global movement (translation): Left: Example initial configuration. Middle: A movement rule applied to the second monomer from the left. It decrements its state to 2 and translates all subsequent monomers by $+\vec{w}$. Right: Another state-3 monomer turns. (d) Blocking: Monomers are not permitted to make movements that would result in a self-intersecting configuration. The monomer m_j is *blocked*: if it were to move, then m_k would move to overlap with m_i .

A turning machine evolves as a continuous-time Markov chain with rotation rules applied asynchronously. This modeling assumption can be used to analyze the running time of the folding process. However, for the purposes of this report, we are only concerned with the order in which the rotation rules are applied. At any moment in time, any monomer with a non-zero state that is not blocked can be the next one to rotate. Therefore, when analyzing the folding process, we assume that the sequence of monomers to which the rules are applied is given by an adversary.

We say that a turning machine successfully folds if eventually all monomers have state 0. Some Turning Machines can evolve into a state where not all monomers have state 0 but no monomer can rotate without the chain self-intersecting. We call these states *permanently blocked*.

Open problems. Kostitsyna et al. [1] have studied the problem of shape formation by a turning machine. Given a desired target shape $S \subset V_\Delta$, specified by a connected subset of nodes of the triangular grid, we consider the problem of creating a turning machine that folds into S . Thus, the goal of the shape formation problem is to assign initial states to the monomers such that the turning machine *always* folds into a desired target shape S (up to a translation / rotation). The paper provides initial results on which classes of shapes can be folded by Turning Machines. It remains open to expand this class of shapes, or prove negative results.

We approached the problem of shape formation from a different angle. Given a turning machine T , we can easily identify the target shape that T encodes. Observe that the differences of the values of initial states of the monomers of T uniquely determine the angles of the chain in the folded state. However, it is challenging to decide whether a given turning machine T always reaches its final folded state, or even if T sometimes reaches its final folded state (i.e., there exists a sequence of rotation rule applications to the monomers that leads to T successfully folding). We know the answer to the question *Do all Turning Machines fold sometimes?*, which is no, there exist Turning Machines that can never be folded. We also know that the question whether a turning machine folds is decidable.

► **Observation 2.** *The state space of a turning machine with $n + 1$ monomers is bounded by $W^n \cdot \text{poly}(n)$, where $W = \max_i \{s_i\} - \min_j \{s_j\}$ is the size of the range of the initial states of the monomers. Hence we can decide whether a turning machine always folds by exploring the (exponential-size) state space.*

We conjecture that both of the questions above are NP-hard. During the seminar we have made partial progress on the way to prove these conjectures.

► **Conjecture 3.** *Given a turning machine T , it is NP-hard to decide whether T always folds.*

► **Conjecture 4.** *Given a turning machine T , it is NP-hard to decide whether T sometimes folds.*

We also considered new tools that may help us to reason about Turning Machines. Below, we provide some observations that we have proven during the seminar. We finish the report with some open questions. We say that a turning machine is *non-negative* if all its monomers have non-negative states. Given two non-negative Turning Machines T and T' , we say that T' *dominates* T if, for each $i \in \{1, 2, \dots, n - 1\}$, it holds that $s'_i \geq s_i$, where s_i and s'_i are the initial states of monomers m_i and m'_i of T and T' , respectively. Here's our first tool.

► **Proposition 3.** *Let T and T' be two non-negative Turning Machines such that T' dominates T . If T' always folds, then T always folds as well.*

Note that the same argument does not apply if we have negative initial states. Rotating a monomer with a negative state (that is, rotating clockwise) in the dominated turning machine corresponds to rotating all the remaining monomers into the counterclockwise direction of the dominating turning machine. We can, however, modify the domination definition as follows.

► **Definition 1.** We say that a turning machine T' *dominates* a turning machine T if, for each $i \in \{1, 2, \dots, n - 1\}$, $s'_i \geq s_i$ if $s_i > 0$ and $s'_i \leq s_i$ if $s_i < 0$ (if $s_i = 0$ then s'_i can take any value).

One of the questions we would like to explore further is whether we can obtain a result similar to Proposition 3 for this new definition of domination. Can the property of domination help us expand the class of Turning Machines for which we can determine whether they always fold, sometimes fold, or never fold? Another question we leave for further investigation is whether there are other “domination” properties that can be formulated for sub- or super-chains that can help answer folding questions.

Conclusions. Kostitsyna et al. [1] initiated the study of the turning machines, a simple model for a molecular folding robot. The Dagstuhl Seminar provided us with a great opportunity to expand the range of research questions that can be studied within this model. We have obtained some preliminary results.

References

- 1 Irina Kostitsyna, Cai Wood, and Damien Woods, *Turning machines: a simple algorithmic model for molecular robotics*, Natural Computing (2022). doi: 10.1007/s11047-022-09880-8
- 2 Damien Woods, *Turning machines! New paper on reconfigurable robotics*, 2020. <https://dna.hamilton.ie/2020-09-16-turning-machines.html>. Accessed 2022-04-09.

4.7 Flips in Higher Order Delaunay triangulations

Anne Driemel (Universität Bonn, Germany, driemel@cs.uni.bonn.de)

Lena Schlipf (Universität Tübingen, Germany, schlipf@informatik.uni.tuebingen.de)

Rodrigo I. Silveira (Universitat Politècnica de Catalunya, Barcelona, rodrigo.silveira@upc.edu)

Jonathan Spreer (University of Sydney, Australia, jonathan.spreer@sydney.edu.au)

License  Creative Commons BY 4.0 International license
 © Rodrigo Silveira, Anne Driemel, Jonathan Spreer, Lena Schlipf

In this working group we consider a class of triangulations of points in \mathbb{R}^2 known as *higher Order Delaunay triangulations* (HODTs) [2]. Given a set of points S , and a parameter k , an *order- k (Delaunay) triangle* is a triangle whose circumcircle contains at most k points from S . An *Order- k (Delaunay) triangulation* is one where all triangles are order- k triangles.

The main topic studied is the flip graph of order- k triangulations. A basic aspect to understand is its connectivity. It is known that it is always connected for $k = 0, 1, 2$, and that it may be not connected for $k \geq 3$, even for points in convex position. Recently, a bit more was understood. In particular, in [1], the following was shown. Let $G(\mathcal{T}_k(S))$ be the flip graph of order- k triangulations of a point set S . Firstly, it was shown that for any k there exists a point set S in convex position where $G(\mathcal{T}_k(S))$ is disconnected. Moreover, $k - 1$ flips are sometimes necessary in order to transform an order- k triangulation of S into another. Secondly, for any order- k triangulation of a point set in convex position there exists some other order- k triangulation at distance at most $k - 1$ in $G(\mathcal{T}_{2k-2}(S))$. Finally, it was also shown that in case $k = 2, 3, 4, 5$, for any order- k triangulation of a point set in general position there exists an order- k triangulation at distance at most $k - 1$ in $G(\mathcal{T}_{2k-2}(S))$. These results imply that the diameter of the flip graph is $O(kn)$.

Discussed Problems

With the long-term goal of understanding the flip graph of order- k HODTs, we focused on some simpler and more concrete questions.

In relation to computation, an interesting question is how fast one can compute the flip distance between two triangulations. Here there are two natural variants of the question: (i) one can go through any triangulation; (ii) you can only go through order- k triangulations. We did not make much progress on this front.

A more combinatorial aspect has to do with understanding the structure of fixed edges, those present in *any* order- k Delaunay triangulation. It is known that already for $k = 2$, the subdivision given by the set of fixed edges can produce polygons that contain holes. During the workshop we have found examples, also for $k = 2$, where holes can be nested. An example with several (non-nested) components for $k = 2$ was also found, but it contains many co-circular points. It is unclear whether the same can be achieved if points are in general position.

Conclusions

We have made some progress in understanding the structure of fixed edges for $k = 2$, but are still far from understanding its flip graph. Results on the flip graph can have direct application to algorithms, either exact or fixed-parameter tractable. For instance, a practical fast algorithm to compute flip distance for $k = 2$ would be very interesting.

References

- 1 Elena Arseneva, Prosenjit Bose, Pilar Cano, and Rodrigo I. Silveira. Flips in higher order Delaunay triangulations. In Yoshiharu Kohayakawa and Flávio Keidi Miyazawa, editors, *LATIN 2020: Theoretical Informatics – 14th Latin American Symposium, São Paulo, Brazil, January 5–8, 2021, Proceedings*, volume 12118 of *Lecture Notes in Computer Science*, pages 223–234. Springer, 2020.
- 2 Joachim Gudmundsson, Mikael Hammar, and Marc J. van Kreveld. Higher order Delaunay triangulations. *Comput. Geom.*, 23(1):85–98, 2002.

5 Open Problems

5.1 Representing Graphs by Polygon Contact in 3D

Will Evans (University of British Columbia, Canada, will@cs.ubc.ca)

Alexander Wolff (Universität Würzburg, Germany)

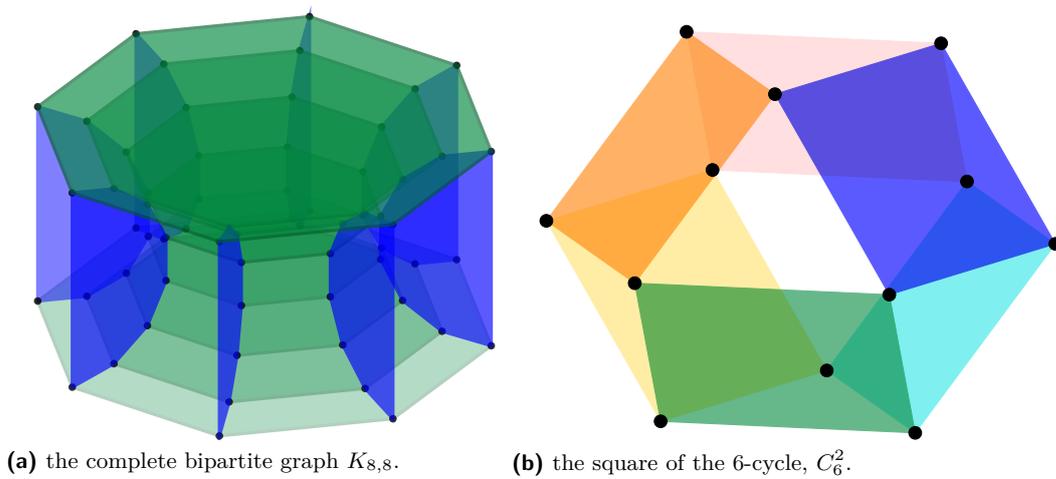
License  Creative Commons BY 4.0 International license
 © Will Evans and Alexander Wolff

We are interested in representing graphs as contact graphs of *convex* polygons in 3D. Adjacency is represented by vertex contacts. Any two polygons must either be disjoint or they can share one vertex; see Figure 26. With others, we [1] showed that any graph admits such a contact representation, but for K_n we need volume $O(n^4n!)$. Let $\text{vol}(n, \Delta)$ be the maximum volume needed for representing any graph with n vertices and vertex degree at most Δ . We know that $\text{vol}(n, 3) = O(n^3)$. The proof was quite tricky, and it seemed to depend a lot on the fact that triangles behave much more nicely than polygons of larger degree. So what is $\text{vol}(n, 4)$? What about lower bounds for $\text{vol}(n, \Delta)$?

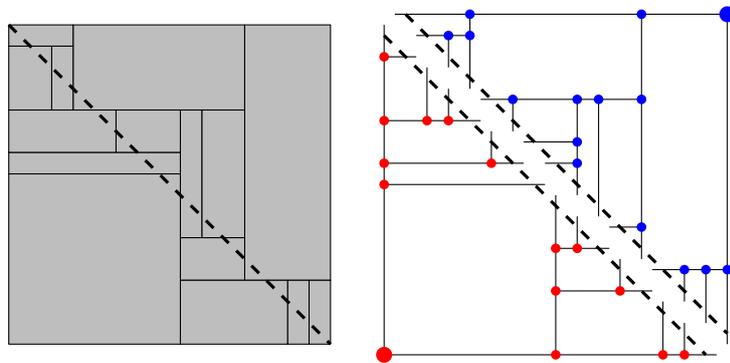
We are also interested in how restrictions on other measures of the graph impact the volume required for its contact representation. In particular, limiting the rectilinear or book thickness of the graph, its k -planarity, or its clique number seem to have the potential to impact this volume.

References

- 1 W. Evans, P. Rzażewski, N. Saeedi, C.-S. Shin, and A. Wolff. Representing graphs and hypergraphs by touching polygons in 3D. In D. Archambault and C. D. Tóth, editors, *Proc. 27th Int. Symp. Graph Drawing & Network Vis. (GD'19)*, volume 11904 of *LNCS*, pages 18–32. Springer, 2019. <http://arxiv.org/abs/1908.08273>. URL: https://doi.org/10.1007/978-3-030-35802-0_2.



■ **Figure 26** Examples of contact representations of graphs using polygons in \mathbb{R}^3 .



■ **Figure 27** Diagonal rectangulation (left) and the corresponding pair of binary trees (right).

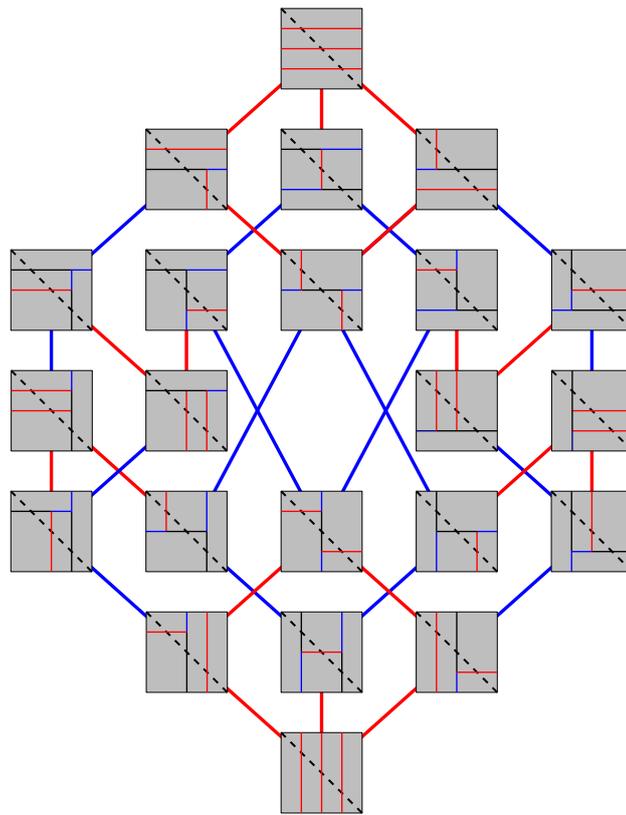
5.2 Diameter of rectangulation flip graph

Torsten Mütze (University of Warwick, UK, torsten.mutze@warwick.ac.uk)

License © Creative Commons BY 4.0 International license
© Torsten Mütze

A *diagonal rectangulation* is a partition of the unit square into finitely many interior-disjoint rectangles, such that no four rectangles meet in a point, and every rectangle intersects the SE-NW-diagonal, drawn dashed in Figure 27. The number of diagonal rectangulations with n rectangles is given by the Baxter numbers, and they are in bijection to $\{2413, 3142\}$ -avoiding permutations, and to pairs of binary trees rooted in the NE and SW corners [2].

We equip the set of all diagonal rectangulations with n rectangles with a *flip* operation that either reverses the orientation of two rectangles whose union is a rectangle, or that rotates one of the arms of a T-join by 90° , which creates a flip graph G_n ; see Figure 28. In this figure, the first type of flips induces the red edges, and the second type of flips induces the blue edges. These flips on rectangulations correspond to rotations in the aforementioned binary trees.



■ **Figure 28** Flip graph G_n of diagonal rectangulations with $n = 4$ rectangles.

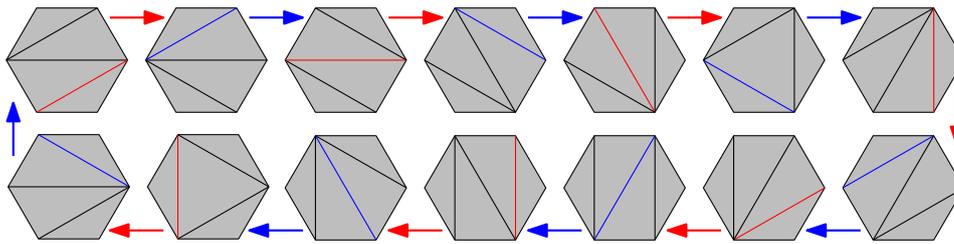
The graph G_n is known to be the cover graph of a lattice, and the skeleton of a polytope [3], very much analogous to the Tamari lattice and the associahedron for triangulations. In this project we are interested in the diameter of G_n , analogous to the famous diameter question for the associahedron [4].

Question: What is the diameter of G_n ?

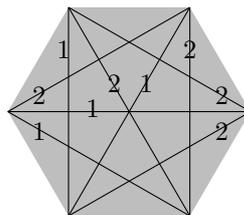
This problem was mentioned by Jean Cardinal at SoCG 2018, and so far only relatively little is known. Ackerman et al. [1] proved an upper bound of $11n$.

References

- 1 Ackerman, E., Allen, M. M., Barequet, G., Löffler, M., Mermelstein, J., Souvaine, D. L., and Tóth, C. D. The flip diameter of rectangulations and convex subdivisions. *Discrete Math. Theor. Comput. Sci.* 18, 3 (2016), Paper No. 4, 17.
- 2 Felsner, S., Fusy, E., Noy, M., and Orden, D. Bijections for Baxter families and related objects. *J. Combin. Theory Ser. A* 118, 3 (2011), 993–1020.
- 3 Law, S., and Reading, N. The Hopf algebra of diagonal rectangulations. *J. Combin. Theory Ser. A* 119, 3 (2012), 788–824.
- 4 Pournin, L. The diameter of associahedra. *Adv. Math.* 259 (2014), 13–42.



■ **Figure 29** Gray code for all 14 triangulations of the 6-gon.



■ **Figure 30** Number of appearances of each edge along the Gray code from Figure 29.

5.3 Balanced triangulation Gray codes

Torsten Mütze (University of Warwick, UK, torsten.mutze@warwick.ac.uk)

License Creative Commons BY 4.0 International license
© Torsten Mütze

We consider all triangulations of a convex n -gon. Two of them differ in a *flip* if they agree in all but 2 triangles. The corresponding flip graph, which has as nodes all triangulations, with edges connecting triangulations that differ in a flip, is the famous associahedron. It is well-known [3, 4] that for any $n \geq 3$, the associahedron has a Hamilton cycle. In other words, there is a Gray code for triangulations, i.e., we can list them cyclically so that any two consecutive triangulations differ in a flip; see Figure 29.

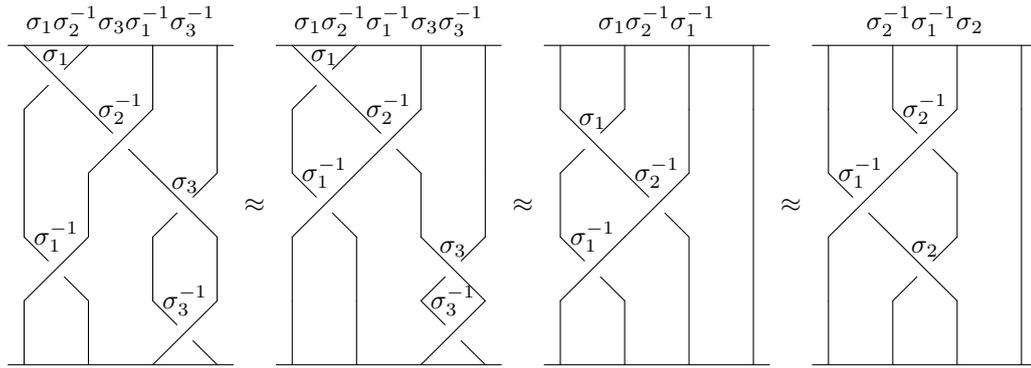
In a recent paper [2], we consider cycles in the associahedron that are *balanced*, i.e., if we count the number of times that each of the $\binom{n}{2} - n$ possible inner edges of the triangulation appears along the cycle, then these counts differ by at most ± 1 ; see Figure 30. Clearly, if an edge appears k times along the cycle, it also has to disappear k times.

Question: Is there a balanced Gray code for triangulations for every $n \geq 3$?

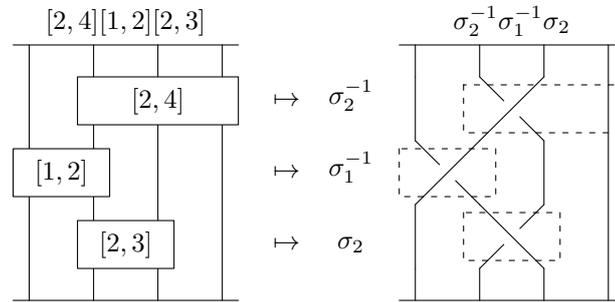
This problem is analogous to the problem of constructing balanced Gray codes for all 2^n binary strings of length n , where one bit is flipped in each step, and each of the n bits should be flipped the same number of times (up to ± 1) [1].

References

- 1 Bhat, G. S., and Savage, C. D. Balanced Gray codes. *Electron. J. Combin.* 3, 1 (1996), Paper 25, 11 pp.
- 2 Felsner, S., Kleist, L., Mütze, T., and Sering, L. Rainbow cycles in flip graphs. *SIAM J. Discrete Math.* 34, 1 (2020), 1–39.
- 3 Hurtado, F., and Noy, M. Graph of triangulations of a convex polygon and tree of triangulations. *Comput. Geom.* 13, 3 (1999), 179–188.
- 4 Lucas, J. M. The rotation graph of binary trees is Hamiltonian. *J. Algorithms* 8, 4 (1987), 503–535.



■ **Figure 31** The braid words $\sigma_1\sigma_2^{-1}\sigma_3\sigma_1^{-1}\sigma_3^{-1}$ and $\sigma_2^{-1}\sigma_1^{-1}\sigma_2$ are isotopic.



■ **Figure 32** The braid word $\sigma_1\sigma_2^{-1}\sigma_3\sigma_1^{-1}\sigma_3^{-1}$ matches the pattern $[2, 4][1, 2][2, 3]$ because the isotopic braid word $\sigma_2^{-1}\sigma_1^{-1}\sigma_2$ matches it directly.

5.4 Pattern-Matching on Braids

Tim Ophelders (Utrecht University & TU Eindhoven, The Netherlands, t.a.e.ophelders@uu.nl)

License © Creative Commons BY 4.0 International license
© Tim Ophelders

The *braid group* on n strands has generators $\sigma_1, \dots, \sigma_{n-1}$ and the following relations.

- $\sigma_i\sigma_j = \sigma_j\sigma_i$ for all $i + 1 < j$.
- $\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$.

Let $\Sigma_{i,j} := \{\sigma_i, \dots, \sigma_{j-1}\} \cup \{\sigma_i^{-1}, \dots, \sigma_{j-1}^{-1}\}$ and $\Sigma_n := \Sigma_{1,n}$. A *braid word* on n strands is a sequence of elements of Σ_n . We say that two braid words are *isotopic* if the corresponding elements of the braid group are equal. See Figure 31 for an example and a geometric interpretation of the braid group.

Let $I_n := \{\{i, j\} \mid i, j \in \{1, \dots, n\}\}$ be the set of intervals with integer endpoints between 1 and n . A *pattern* on n strands is a sequence of elements of $\Sigma_n \sqcup I_n$.

A braid word *directly matches* a pattern p if it can be obtained from p by substituting each element $[i, j]$ by some sequence of elements of $\Sigma_{i,j}$ (different occurrences of $[i, j]$ may be substituted by different such sequences). A braid word *matches* a pattern p if it is isotopic to a braid word that directly matches p , see Figure 32.

Question 1. Given a braid b and a pattern p , is testing whether b matches p decidable?

Call a pattern *pure* if all of its elements are drawn from I_n , i.e. none are drawn from Σ_n .

Subquestion 1.1. Given a braid b and a pure pattern p , is testing whether b matches p decidable?

Participants

- Florestan Brunck
IST Austria –
Klosterneuburg, AT
- Kevin Buchin
TU Eindhoven, NL
- Maïke Buchin
Ruhr-Universität Bochum, DE
- Benjamin Burton
The University of Queensland –
Brisbane, AU
- Jean Cardinal
ULB – Brussels, BE
- Éric Colin de Verdière
CNRS & LIGM –
Marne-la-Vallée
- Arnaud de Mesmay
University Paris-Est –
Marne-la-Vallée, FR
- Linda Kleist
TU Braunschweig, DE
- Boris Klemz
Universität Würzburg, DE
- Irina Kostitsyna
TU Eindhoven, NL
- Francis Lazarus
CNRS – Grenoble, FR
- Maarten Löffler
Utrecht University, NL
- Alexander Neuhaus
Ruhr-Universität Bochum, DE
- Tim Ophelders
Utrecht University, NL
- Alexander Wolff
Universität Würzburg, DE



Remote Participants

- Mark Bell
Hampshire, GB
- Hsien-Chih Chang
Dartmouth College –
Hanover, US
- Vincent Delecroix
University of Bordeaux, FR
- Anne Driemel
Universität Bonn, DE
- Nathan Dunfield
University of Illinois –
Urbana Champaign, US
- William Evans
University of British Columbia –
Vancouver, CA
- Brittany Terese Fasy
Montana State University –
Bozeman, US
- Fabrizio Frati
University of Rome III, IT
- Niloufar Fuladi
University Paris-Est –
Marne-la-Vallée, FR
- Elise Goujard
University of Bordeaux, FR
- Luke Jeffreys
University of Bristol, GB
- Anna Lubiw
University of Waterloo, CA
- Torsten Mütze
University of Warwick –
Coventry, GB
- Hugo Parlier
University of Luxembourg, LU
- Lionel Pournin
Université Paris Nord –
Villetaneuse, FR
- Jessica S. Purcell
Monash University –
Clayton, AU
- Saul Schleimer
University of Warwick –
Coventry, GB
- Lena Schlipf
Universität Tübingen, DE
- Eric Sedgwick
DePaul University – Chicago, US
- Rodrigo I. Silveira
UPC Barcelona Tech, ES

- Jonathan Spreer
University of Sydney, AU

- Stephan Tillmann
University of Sydney, AU

- Uli Wagner
IST Austria –
Klosterneuburg, AT

- Carola Wenk
Tulane University –
New Orleans, US

- Erin Moriarty Wolf Chambers
St. Louis University, US

New Perspectives in Symbolic Computation and Satisfiability Checking

Erika Abraham^{*1}, James H. Davenport^{*2}, Matthew England^{*3}, and Alberto Griggio^{*4}

1 RWTH Aachen University, DE. abraham@informatik.rwth-aachen.de

2 University of Bath, GB. J.H.Davenport@bath.ac.uk

3 Coventry University, GB. Matthew.England@coventry.ac.uk

4 Fondazione Bruno Kessler – Trento, IT. griggio@fbk.eu

Abstract

Dagstuhl Seminar 22072 gathered researchers from Symbolic Computation and Satisfiability Checking. These communities have independent histories but worked together in recent years (e.g. Dagstuhl Seminar 15471 and the EU SC-Square Project). We seek to tackle problems which are in the interest of both communities, and require the expertise of both to overcome.

Seminar February 13–18, 2022 – <http://www.dagstuhl.de/22072>

2012 ACM Subject Classification Computing methodologies → Symbolic and algebraic manipulation; Theory of computation → Automated reasoning

Keywords and phrases computer algebra systems, SMT Solvers, verification

Digital Object Identifier 10.4230/DagRep.12.2.67

1 Executive Summary

Matthew England (Coventry University, GB)

Erika Abraham (RWTH Aachen University, DE)

James H. Davenport (University of Bath, GB)

Alberto Griggio (Fondazione Bruno Kessler, IT)

License © Creative Commons BY 4.0 International license

© Matthew England, Erika Abraham, James H. Davenport, and Alberto Griggio

Introduction

Symbolic Computation refers to algorithms for computers to perform symbolic mathematics, usually implemented in Computer Algebra Systems (CASs). Satisfiability Checking refers to algorithms to efficiently check the satisfiability of a logical statement, developed originally for the Boolean domain and implemented in SAT solvers, but now extended to a wide variety of different theories in satisfiability modulo theories (SMT) solvers. This Dagstuhl Seminar is on Symbolic Computation and Satisfiability Checking, with the emphasis on the “and” to indicate the scope is strictly work of interest to both communities.

Traditionally, the two communities have been largely disjoint and unaware of the achievements of one another, despite there being strong reasons for them to discuss and collaborate, since they share many central interests. Many of the theories tackled by SMT have been traditionally studied within Symbolic Computation; while in the opposite direction, the integration of SAT solvers into computer algebra systems can allow more powerful logical reasoning and inspire new algorithmic approaches in computer algebra.

* Editor / Organizer



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

New Perspectives in Symbolic Computation and Satisfiability Checking, *Dagstuhl Reports*, Vol. 12, Issue 2, pp. 67–86

Editors: Erika Abraham, James H. Davenport, Matthew England, and Alberto Griggio



DAGSTUHL
REPORTS Dagstuhl Reports

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

Recent History

The first global meeting dedicated to both symbolic computation and satisfiability checking was Dagstuhl Seminar 15471 (Symbolic Computation and Satisfiability Checking) [1] which took place in November 2015. This was followed soon after by EU Horizon 2020 Grant 712689 which ran from 2016-2018. The aim of that project was to bridge the gap between the communities to produce individuals who can combine the knowledge and techniques of both fields to resolve problems currently beyond the scope of either [2]. The project funded new collaborations, new tool integrations, proposals on extensions to the SMT-LIB language standards, new collections of benchmarks, two summer schools (in 2017 and 2018) and the SC-Square Workshop Series.

The Workshop Series (<http://www.sc-square.org/workshops.html>) has taken place annually for six years, with two further editions already planned:

2016 Timișoara, Romania (as part of SYNASC 2016).

2017 Kaiserslautern, Germany (alongside ISSAC 2017).

2018 Oxford, UK (as part of FLoC 2018).

2019 Bern, Switzerland (as part of SIAM AG19)

2020 Paris, France (online) (alongside IJCAR 2020)

2021 Texas, USA (online) (as part of SIAM AG21)

2022 Haifa, Israel (as part of FLoC 2022)

2023 Tromsø Norway (alongside ISSAC 2023)

It takes place as part of, or alongside, established conferences (alternating between computational algebra and logic). Each year there are two chairs, one from each community.

In 2020 a special issue of the Journal of Symbolic Computation was published, on the theme of SC-Square [3]. A further special issue is in development.

Motivation for new Seminar

The seminar call defined its scope with these research questions.

Decision Procedures: How to efficiently leverage CAS for SMT over hard arithmetical theories? How to exploit conflict-driven learning and non-chronological backtracking in symbolic computation algorithms? How can CAS and SMT be combined to reason about bit-precise machine (i.e. floating point) arithmetic?

Abstraction and Linearization: How can abstraction techniques commonly adopted in SMT be exploited in symbolic computation? How to leverage techniques in CASs for iterative abstraction refinement in SMT?

Optimization: Can SMT and symbolic computation be combined for successfully attacking non-linear optimization problems? Can new optimization techniques be leveraged for heuristic choices in solvers?

Machine Learning: What are the common challenges and opportunities on the use of Machine Learning (ML) for heuristic choices in algorithms? How best to define problem features for classic ML? How best to encode formulae for deep ML? How to develop good datasets for ML? Tool development: How to share data structures, low-level libraries, input formats and interaction pipelines for more effective development of robust, mature and interoperable symbolic reasoning tools?

Application Problem Encoding: How best encode high-level application problems to be more amenable to symbolic reasoning? How to provide more expressive problem definition languages which can still be handled efficiently? How to automate problem encoding?

Seminar Overview

The seminar was organised into eight sessions by broad topic (with some exceptions to allow for online participants). We invited three extended tutorials on key topics of interest to the seminar: Ahmed Irfan spoke on the incremental linearization techniques developed for MathSAT to tackle non-linear problems, including ones involving transcendental functions; Haniel Barbosa and Gereon Kremer described the new work on proof certificates in CVC5, and the possibilities for extensions into non-linear real arithmetic; and Curtis Bright spoke on isomorphism free exhaustive generation techniques which used a combination of computer algebra and SAT solvers. Other talks were contributed by seminar participants.

Upcoming Development

Erika Ábrahám, Chris Brown, James Davenport, Pascal Fontaine and Thomas Sturm are the joint editors of a Journal of Symbolic Computation Special Issue on the topic of “Symbolic Computation and Satisfiability Checking”. Contributions coming out of this workshop would be especially welcomed. The timeline is given below.

31 March Submissions Open

31 August Submissions Close

(early notification to abraham@cs.rwth-aachen.de is welcomed.)

31 December Authors notified

3–6 months Articles published

Special issues are now “virtual” and so the articles appear online as ready.

Acknowledgements

The present seminar was originally scheduled for November 2020 but was re-organised for February 2022 following the COVID19 pandemic. It was held in a hybrid format, with participants both in person and online. The organisers thank all participants for their contribution, and extend a special thanks to the team at Schloss Dagstuhl for their excellent organisation and support, particularly in the difficult circumstance which the pandemic imposed.

References

- 1 E. Ábrahám and P. Fontaine and T. Sturm and D. Wang. *Symbolic Computation and Satisfiability Checking (Dagstuhl Seminar 15471)*. Dagstuhl Reports 5(11):71-89. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2016. URL: <http://drops.dagstuhl.de/opus/volltexte/2016/5765>
- 2 E. Ábrahám, J. Abbott, B. Becker, A.M. Bigatti, M. Brain, B. Buchberger, A. Cimatti, J.H. Davenport, M. England, P. Fontaine, S. Forrest, A. Griggio, D. Kroening, W.M. Seiler, and T. Sturm. *SC2: Satisfiability checking meets symbolic computation*. In M. Kohlhase, M. Johansson, B. Miller, L. de Moura, and F. Tompa, editors, Intelligent Computer Mathematics: Proceedings CICM 2016, volume 9791 of Lecture Notes in Computer Science, pages 28–43. Springer International Publishing, 2016. URL: https://doi.org/10.1007/978-3-319-42547-4_3
- 3 J.H. Davenport, M. England, A. Griggio, T. Sturm, and C. Tinelli. *Symbolic computation and satisfiability checking: Editorial*. Journal of Symbolic Computation, 100:1–10, 2020. URL: <https://doi.org/10.1016/j.jsc.2019.07.017>

2 Table of Contents

Executive Summary

Matthew England, Erika Abraham, James H. Davenport, and Alberto Griggio . . . 67

Overview of Talks

Better SMT Proofs for Certifying Compliance <i>Haniel Barbosa</i>	72
Comprehensive Groebner Systems (with CoCoA) <i>Anna Maria Bigatti</i>	72
Sinful Behaviour <i>Martin Brain</i>	73
Isomorph-Free Exhaustive Generation in SAT Solving <i>Curtis Bright</i>	73
What SAT/SMT Has Taught Me <i>Christopher W. Brown</i>	74
Towards Scalable Computation of Semi-Algebraic Systems Driven by Applications <i>Changbo Chen</i>	74
SMT-based analysis of Switching Kirchhoff networks <i>Alessandro Cimatti</i>	75
Varieties of Doubly-Exponential behaviour in Quantifier Elimination and Cylindrical Algebraic Decomposition <i>James H. Davenport</i>	75
Using Machine Learning in SC ² <i>Tereso del Rio</i>	76
A Note on SMT-LIB NRA <i>Bruno Dutertre</i>	76
SMT Solving in the Cloud <i>Bruno Dutertre</i>	76
Observations and Questions on the SMT-LIB <i>Matthew England</i>	76
Scalable Optimal Deployment in the Cloud of Component-based Applications using Constraint Programming, Optimization Modulo Theory, Mathematical Programming and Symmetry Breaking <i>Madalina Erascu</i>	77
Recent Logic Improvements in Maple <i>Jürgen Gerhard</i>	78
A Tutorial on Incremental Linearization <i>Ahmed Irfan</i>	78
Heuristic Techniques for Natural Style Proofs in Elementary Analysis <i>Tudor Jebelean</i>	78
Guessing with Little Data <i>Manuel Kauers</i>	79

Satisfiable Algebraic Circuit Verification <i>Daniela Kaufmann</i>	79
20+ Years of Legendre Pairs <i>Ilias S. Kotsireas</i>	80
Formal Proofs for Cylindrical Algebraic Coverings <i>Gereon Kremer</i>	80
Proving Satisfiability in NTA via Unconstrained Optimization and the Topological Degree Test <i>Enrico Lipparini</i>	80
Implementation and Application of Chordality Preserving Top-Down Algorithms for Triangular Decomposition <i>Chenqi Mou</i>	81
Dynamical System Verification using Abstractions and Non-Linear Arithmetic <i>Sergio Mover</i>	81
Levelwise Construction of a Single Cylindrical Algebraic Cell <i>Jasper Nalbach</i>	82
Proof theory and computational algebra <i>Thomas Powell</i>	82
Looking Backward and Forward <i>Stefan Ratschan</i>	83
Computational Limits of Using CAD and SMT Solvers in Logical Analysis of Regulatory Networks <i>AmirHosein Sadeghimanesh</i>	83
Expansion-Based QBF Solving for QBF <i>Martina Seidl</i>	83
Decidable Logics with Arithmetic and Uninterpreted Symbols for SMT <i>Baptiste Vergain</i>	84
Participants	85
Remote Participants	85

3 Overview of Talks

3.1 Better SMT Proofs for Certifying Compliance

Haniel Barbosa (Federal University of Minas Gerais-Belo Horizonte, BR)

License © Creative Commons BY 4.0 International license
© Haniel Barbosa

Joint work of Haniel Barbosa, Andrew Reynolds, Gereon Kremer, Hanna Lachnitt, Aina Niemetz, Andres Nötzli, Alex Ozdemir, Mathias Preiner, Arjun Viswanathan, Scott Viteri, Yoni Zohar, Cesare Tinelli, Clark Barrett

SMT solvers can be hard to trust, since it generally means assuming their large and complex codebases do not contain bugs that lead to wrong results. Machine-checkable certificates, via proofs of the logical reasoning the solver has performed, address this issue by decoupling confidence in the results from the solver’s implementation.

Despite previous work, in several SMT solvers, to produce and check proofs, users still have to choose among solvers that may not produce fine-grained proofs, may not produce proofs for some of their crucial-for-efficiency components, or have proofs that are checkable only as part of a specific proof assistant.

To facilitate the use of SMT proof certificates, the *cvc5* developers team has completely redesigned its proof-production infrastructure, aiming for a sufficiently general and extensible infrastructure to allow: the generation of coarse- and fine-grained proofs for all parts of the solver, particularly for previously unsupported components such as the rewriter and the strings subsolver; and the printing of proofs in different formats to enable the use of different proof checkers. Specifically, we are working on producing *cvc5* proofs for LFSC, Isabelle/HOL, Lean4, and Coq, while also creating proof calculi for previously unsupported SMT-LIB theories in these settings. While the project is still ongoing, we will report on significant progress on all of these fronts.

3.2 Comprehensive Groebner Systems (with CoCoA)

Anna Maria Bigatti (University of Genova, IT)

License © Creative Commons BY 4.0 International license
© Anna Maria Bigatti

Joint work of Elisa Palezzato, Michele Torielli, Anna Maria Bigatti

A comprehensive Groebner system (CGS) is a collection of Groebner bases and algebraic sets describing a parametric polynomial system. For a specialization of parameters, a Groebner basis of the specialized ideal can be immediately recovered from a branch of the associated CGS. This property makes its computation attractive in applications where a family of problems can be formulated as a parametric polynomial system.

The first algorithms [7, 8, 3, 5] required computations in polynomial ring over a coefficient field of rational functions, $K(A)[X]$, where A are the parameters, and X the actual indeterminates, together with delicate handling of the case distinctions over the parameters. This last fact makes these algorithms hard to implement in computer algebra systems.

In 2006 Suzuki-Sato [6] introduced a new approach, further improved in [2, 4], which just needs the computation of Groebner bases in $K[A, X]$, so that that can be easily implemented in any computer algebra system.

In this seminar we present a short history of this latter method, we describe our implementation in CoCoA [1] of their algorithms and of our iterative alternative.

References

- 1 J. Abbott, A. M. Bigatti, and L. Robbiano. *CoCoA: a system for doing Computations in Commutative Algebra*. URL: <https://cocoa.dima.unige.it>
- 2 D. Kapur, Y. Sun, D. Wang. *A new algorithm for computing comprehensive Groebner systems*. In: Watt, S. (ed.) International Symposium on Symbolic and Algebraic Computation, pp. 29–36. ACM Press, 2010.
- 3 M. Manubens, and A. Montes. *Improving the DISPGB algorithm using the discriminant ideal*. J. Symb. Comp. 41(11), 1245–1263, 2006.
- 4 K. Nabeshima. *Stability Conditions of Monomial Bases and Comprehensive Groebner Systems*. CASC’12: Proceedings of the 14th international conference on Computer Algebra in Scientific Computing September 2012, 248–259, 2012.
- 5 A. Suzuki, and Y. Sato. *An Alternative approach to Comprehensive Gröbner Bases*. J. Symb. Comp. 36(3-4), 649–667, 2003.
- 6 A. Suzuki, and Y. Sato. *A Simple Algorithm to Compute Comprehensive Gröbner Bases Using Gröbner Bases*. ISSAC ’06: Proceedings of the 2006 International Symposium on Symbolic and Algebraic Computation, 326–331, 2006.
- 7 V. Weispfenning. *Comprehensive Gröbner bases*, J. Symb. Comp. 14(1), 1–29, 1992.
- 8 V. Weispfenning. *Canonical Comprehensive Gröbner bases*, J. Symb. Comp. 36, 669–683, 2003.

3.3 Sinful Behaviour

Martin Brain (City – University of London, GB)

License © Creative Commons BY 4.0 International license
© Martin Brain

This talk presents an exposition on the nature of sin, and other trigonometric functions and the consequences for the verification of software that use standard mathematical libraries. The counter-intuitive properties of even the ideal implementation are compounded by the underspecified and ambiguous nature of current implementations. We discuss the difficulties of implementing and verifying trigonometric functions and finish with some challenges for how computer algebra might help.

3.4 Isomorph-Free Exhaustive Generation in SAT Solving

Curtis Bright (University of Windsor, CA)

License © Creative Commons BY 4.0 International license
© Curtis Bright

This tutorial will provide an introduction to methods for exhaustively generating combinatorial objects while avoiding isomorphic copies of those objects. The “recorded objects” and “orderly generation” methods from the symbolic computation literature will be described and contrasted with the “symmetry breaking” approach from the satisfiability literature. A method of combining isomorph-free exhaustive generation with a SAT solver will be applied to two problems and shown to improve the performance of the solver by orders of magnitude. It will be argued there is great potential to be unlocked by exploiting both the symbolic and SAT approaches simultaneously.

References

- 1 Curtis Bright, Kevin K. H. Cheung, Brett Stevens, Ilias Kotsireas, and Vijay Ganesh. A SAT-based resolution of Lam’s problem. In *Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence*, pages 3669–3676, 2021.
- 2 Michael Codish, Alice Miller, Patrick Prosser, and Peter J Stuckey. Constraints for symmetry breaking in graph representation. *Constraints*, 24(1):1–24, 2019.
- 3 I. A. Faradžev. Constructive enumeration of combinatorial objects. In *Problèmes combinatoires et théorie des graphes*, pages 131–135, 1978.
- 4 Tommi Junttila, Matti Karppa, Petteri Kaski, and Jukka Kohonen. An adaptive prefix-assignment technique for symmetry reduction. *Journal of Symbolic Computation*, 99:21–49, 2020.
- 5 Brendan D. McKay and Adolfo Piperno. Practical graph isomorphism, II. *Journal of Symbolic Computation*, 60:94–112, 2014.
- 6 Ronald C. Read. Every one a winner or how to avoid isomorphism search when cataloguing combinatorial configurations. In B. Alspach, P. Hell, and D.J. Miller, editors, *Algorithmic Aspects of Combinatorics*, volume 2 of *Annals of Discrete Mathematics*, pages 107–120. Elsevier, 1978.

3.5 What SAT/SMT Has Taught Me

Christopher W. Brown (U.S. Naval Academy – Annapolis, US)

License  Creative Commons BY 4.0 International license
© Christopher W. Brown

SAT/SMT solvers follow a very different paradigm than algorithms from the computer algebra community. In particular, there is an emphasis on bottom-up, conflict-driven approaches. I will look at how this new (to us) paradigm has changed real polynomial constraint solving so far, and where it might take us in the future.

3.6 Towards Scalable Computation of Semi-Algebraic Systems Driven by Applications

Changbo Chen (Chinese Academy of Sciences – Chongqing, CN)

License  Creative Commons BY 4.0 International license
© Changbo Chen

Real applications require to solve semi-algebraic systems with number of variables ranging from a few to dozens, hundreds, thousands, or even millions or billions. It is a great challenge to make the nonlinear solvers in computer algebra scalable to solve large systems. In this talk, we first review our recent works on increasing the scale of semi-algebraic systems solving by exploiting structures from applications. Then we discuss some possible ways to make RC-CAD, which is an approach for computing CAD based on regular chains, scalable.

3.7 SMT-based analysis of Switching Kirchhoff networks

Alessandro Cimatti (Fondazione Bruno Kessler, IT)

License  Creative Commons BY 4.0 International license
© Alessandro Cimatti

Complex dynamical systems require automated analysis techniques and tools. In this talk I focus on a challenging class of systems that can be described as switching multi-domain Kirchhoff networks (SMDKN), where the global behaviour result from combining the behaviour of components by way of conservation laws. In the first part of the talk, I present a comprehensive, long-term picture where the problems of validating and reformulating an SMDKN description in terms of Differential-Algebraic Equations (DAE) into hybrid automata equipped with Ordinary Differential Equations (DAE). The approach is based on modern satisfiability and verification modulo theory (SMT and VMT) techniques over nonlinear and transcendental functions (NTA) [3]. In the second part of the talk, I discuss a practical subcase, proposing an engineering-oriented approach for the formal analysis of Relay-based Railways Interlocking Systems, developed in the context of an industrial collaboration with the Italian Railway Network company [2, 1].

References

- 1 Arturo Amendola, Anna Becchi, Roberto Cavada, Alessandro Cimatti, Andrea Ferrando, Lorenzo Pilati, Giuseppe Scaglione, Alberto Tacchella, and Marco Zamoni. *NORMA: a tool for the analysis of Relay-based Railway Interlocking Systems*. Proc. TACAS '22, Lecture Notes in Computer Science, 13243, pages 125–142. Springer, 2022.
- 2 Roberto Cavada, Alessandro Cimatti, Sergio Mover, Mirko Sessa, Giuseppe Cadavero, Giuseppe Scaglione. *Analysis of Relay Interlocking Systems via SMT-based Model Checking of Switched Multi-Domain Kirchhoff Networks*, Proc. FMCAD '18, pages 1–9. IEEE, 2018.
- 3 Alessandro Cimatti, Sergio Mover, Mirko Sessa. *SMT-based analysis of switching multi-domain linear Kirchhoff networks*, Proc. FMCAD '17, pages 188–195. IEEE, 2017.

3.8 Varieties of Doubly-Exponential behaviour in Quantifier Elimination and Cylindrical Algebraic Decomposition

James H. Davenport (University of Bath, GB)

License  Creative Commons BY 4.0 International license
© James H. Davenport

It is 45 years since Davenport and Heintz drafted “Real Quantifier Elimination is Doubly Exponential”. In the natural representation, both the number of polynomials and the degree are doubly exponential in the number of variables. This talk looks at the varieties of doubly exponential behaviour for various algorithms. These can be growing doubly exponentially in both the number and degree of the polynomials involved [1]. This is inherent in resultant-based processes. We ask whether this is inherent in other methods, such as Virtual Term Substitution and Comprehensive Gröbner Systems. It *may* not be, but this is a significant research question.

References

- 1 M. England and J.H. Davenport. *The Complexity of Cylindrical Algebraic Decomposition with Respect to Polynomial Degree*. Proceedings CASC 2016, V.P. Gerdt, W. Koepf and W.M. Seiler (eds), LNCS 9890, pages 172–192. Springer, 2016. URL: https://doi.org/10.1007/978-3-319-45641-6_12

3.9 Using Machine Learning in SC^2

Tereso del Rio (Coventry University, GB)

License  Creative Commons BY 4.0 International license
© Tereso del Rio

This talk exposes many possible uses of Machine Learning (ML) in the context of SC^2 , and how this approach differs from human-made heuristics. Different ML paradigms are presented and it is discussed how symbolic data could be encoded. This talk also intends to motivate a discussion about which datasets should be used for training and testing ML models.

3.10 A Note on SMT-LIB NRA

Bruno Dutertre (Amazon – Cupertino, US)

License  Creative Commons BY 4.0 International license
© Bruno Dutertre

The talk discusses some consequences of the semantics of division by zero adopted by the SMT-LIB standard.

3.11 SMT Solving in the Cloud

Bruno Dutertre (Amazon – Cupertino, US)

License  Creative Commons BY 4.0 International license
© Bruno Dutertre

We will discuss challenges and opportunities in large-scale parallelism for SMT solving.

3.12 Observations and Questions on the SMT-LIB

Matthew England (Coventry University, GB)

License  Creative Commons BY 4.0 International license
© Matthew England

In this seminar we start by giving an overview on the SMT-LIB initiative [1]: an incredible community driven and volunteer run resource. The SMT-LIB provides rigorous descriptions of background theories, develops a common input language for AMT-solvers, and maintains a large and growing library of benchmarks. The author, whose background is Symbolic Computation, was impressed by the SMT-LIB as there was no comparable resource for computer algebra. The paper [2] offers a good explanation of benchmarking in SAT/SMT to the newcomer from computer algebra.

However, he notes that some care must be taken when using it for the training of machine learning models, or benchmarking more generally. Focussing on the non-linear arithmetic section, we note for example: how very different conclusions can be drawn by the exclusion/inclusion of the MetiTarski dataset (which is an order of magnitude larger than all other datasets in this section); how a large part of these benchmarks can be solved without

theory calls, and the vast majority without recourse to complete NRA algorithms; and how some basic simplification routines can change many problems from one theory to a simpler one.

Such issues raise questions about the best way to use this resource and how the resource itself may be improved. The author suggests that a follow up paper to [2] may be required to outline best practice in use of these benchmarks.

References

- 1 C. Barrett, P. Fontaine, and C. Tinelli. *The Satisfiability Modulo Theories Library (SMT-LIB)*. URL: <https://smtlib.cs.uiowa.edu/>
- 2 M. Brain, J.H. Davenport, and A. Griggio. *Benchmarking solvers, SAT-style*. In M. England and V. Ganesh, editors, *Proceedings of the 2nd International Workshop on Satisfiability Checking and Symbolic Computation (SC² 2017)*, number 1974 in *CEUR Workshop Proceedings*, 2017. URL: <http://ceur-ws.org/Vol-1974/>

3.13 Scalable Optimal Deployment in the Cloud of Component-based Applications using Constraint Programming, Optimization Modulo Theory, Mathematical Programming and Symmetry Breaking

Madalina Erascu (West University of Timisoara, RO)

License © Creative Commons BY 4.0 International license
© Madalina Erascu

Joint work of Erascu, Madalina; Micota, Flavia; Zaharie, Daniela

Main reference Madalina Erascu, Flavia Micota, Daniela Zaharie: “Scalable optimal deployment in the cloud of component-based applications using optimization modulo theory, mathematical programming and symmetry breaking”, *J. Log. Algebraic Methods Program.*, Vol. 121, p. 100664, 2021.

URL <http://dx.doi.org/10.1016/j.jlamp.2021.100664>

Automated deployment of component-based applications in the Cloud consists in the allocation of virtual machines (VMs) offers from various Cloud Providers such that the constraints induced by the interactions between components and by the components hardware/software requirements are satisfied and the performance objectives are optimized (e.g. costs are minimized). It can be formulated as a constraint optimization problem, hence, in principle, the optimization can be carried out automatically. In the case the set of VM offers is large (several hundreds), the computational requirement is huge, making the automatic optimization practically impossible with the current general optimization modulo theory (OMT) and mathematical programming (MP) tools. We overcame the difficulty by methodologically analyzing the particularities of the problem with the aim of identifying search space reduction methods. These are methods exploiting: (1) the symmetries of the general Cloud deployment problem, (2) the graph representation associated to the structural constraints specific to each particular application, and (3) their combination. An extensive experimental analysis has been conducted on four classes of real-world problems, using six symmetry breaking strategies and two types of optimization solvers. As a result, the combination of a variable reduction strategy with a column-wise symmetry breaker leads to a scalable deployment solution, when OMT is used to solve the resulting optimization problem.

3.14 Recent Logic Improvements in Maple

Jürgen Gerhard (Maplesoft – Waterloo, CA)

License  Creative Commons BY 4.0 International license
© Jürgen Gerhard

The presentation will focus on some of the new features and improvements in recent releases of Maple in the areas of computational logic, which are of particular interest to the participants of this seminar. Topics included are Boolean logic and SAT solvers, SMTLIB, polyhedral sets, and nonlinear real arithmetic.

3.15 A Tutorial on Incremental Linearization

Ahmed Irfan (Amazon – Cupertino, US)

License  Creative Commons BY 4.0 International license
© Ahmed Irfan

Incremental linearization is a simple and practical approach to decide the satisfiability of first-order formulas containing nonlinear arithmetic and transcendental functions. The key idea is to use the abstraction-refinement method by abstracting nonlinear multiplication and transcendental functions as uninterpreted functions, allowing us to leverage efficient methods for linear arithmetic. The abstraction is refined iteratively by axiomatizing the uninterpreted functions by upper- and lower-bounding piecewise linear constraints.

In this tutorial talk, I will walk through the key ideas of the technique and give details for practical consideration. I will also touch upon its recent developments and related open research directions.

3.16 Heuristic Techniques for Natural Style Proofs in Elementary Analysis

Tudor Jebelean (Universität Linz, AT)

License  Creative Commons BY 4.0 International license
© Tudor Jebelean

Automatic construction of proofs in rich theories (like e. g. elementary analysis) is difficult because the purely logic approach cannot efficiently handle the relatively large number of necessary properties and especially the algorithms based on them. Therefore it looks promising to combine logic with domain specific methods, which is basically equivalent to SMT solving. Our goal goes a little beyond, that is we are aiming at producing automatically proofs that can be easily understood by human readers. For this purpose we identified several heuristic techniques:

- the S-decomposition method for formulae with alternating quantifiers;
- quantifier elimination by cylindrical algebraic decomposition;
- analysis of terms behavior in zero;
- bounding the ϵ -bounds;
- semantic simplification of expressions involving absolute value;

- polynomial arithmetic and solving;
- usage of equal arguments to arbitrary functions; and
- reordering of proof steps in order to insure the admissibility of solutions to meta-variables.

These techniques allow to produce natural-style proofs for many interesting examples [1], like convergence of sum and product of sequences, continuity of sum, product and compositions of functions, etc. As proving in the theory of reals is akin to satisfiability modulo this theory, one expects that these heuristic techniques may inspire more general methods for SMT solving.

References

- 1 T. Jebelean. *A Heuristic Prover for Elementary Analysis in Theorema*. Intelligent Computer Mathematics (CICM 2021), Lecture Notes in Computer Science 12833:130–134. Springer, 2021. URL: https://doi.org/10.1007/978-3-030-81097-9_10

3.17 Guessing with Little Data

Manuel Kauers

License  Creative Commons BY 4.0 International license
© Manuel Kauers

Joint work of Manuel Kauers, Christoph Koutschan

Main reference Manuel Kauers, Christoph Koutschan: “Guessing with Little Data”, CoRR, Vol. abs/2202.07966, 2022.

URL <https://arxiv.org/abs/2202.07966>

Reconstructing a hypothetical recurrence equation from the first terms of an infinite sequence is a classical and well-known technique in experimental mathematics. We say that such an equation is found by guessing. The success of this approach depends on how big the recurrence equation is and how many terms of the sequence are known. The bigger the equation, the more terms are needed to reliably find it.

There are sequences for which it is difficult to generate a lot of terms, and they may satisfy recurrence equations that out of reach of the common guessing algorithm. We present a variation of the guessing algorithm which can succeed with significantly fewer input terms.

This is joint work with Christoph Koutschan [1].

References

- 1 Manuel Kauers and Christoph Koutschan. Guessing with little data. Arxiv 2202.07966.

3.18 Satisfiable Algebraic Circuit Verification

Daniela Kaufmann (Johannes Kepler Universität Linz, AT)

License  Creative Commons BY 4.0 International license
© Daniela Kaufmann

Joint work of Daniela Kaufmann, Armin Biere, Jakob Nordström, Paul Beame

Although algebraic reasoning is one of the most successful methods for verifying gate-level integer multipliers, it has limitations with particular components, necessitating the use of SAT solvers in addition. As a result, proofs in two different formats are required for validation certifications. The validation results can only be trusted up to compositional reasoning, because approaches to unifying certificates are not scalable. The use of dual variables in

the algebraic encoding and replicating SAT-based notions in polynomial reasoning, on the other hand, eliminates the need for SAT solvers in the verification flow, resulting in a single, uniform proof certificate. In this session, I will discuss open issues in incorporating dual variables into algebraic reasoning.

3.19 20+ Years of Legendre Pairs

Ilias S. Kotsireas (Wilfrid Laurier University – Waterloo, CA)

License  Creative Commons BY 4.0 International license
© Ilias S. Kotsireas

Legendre pairs were introduced in 2001 by Seberry and her students, as a means to construct Hadamard matrices via a two-circulant core construction. A Legendre pair consists of two sequences of the same odd length ℓ , with elements from $-1, +1$, such that their respective autocorrelation coefficients sum to -2 , or (equivalently) their respective power spectral density coefficients sum to $2\ell + 2$. Legendre pairs of every odd prime length exist, via a simple construction using the Legendre symbol. We will review known constructions for Legendre pairs. We will discuss various results on Legendre pairs during the past 20 years, including the concept of compression, introduced in a joint paper with Djokovic, as well as the computational state-of-the-art of the search for Legendre pairs. In particular, we recently contributed the only known Legendre pair of length $\ell = 77$ in a joint paper with Turner/Bulutoglu/Geyer. In addition, we recently contributed in a joint paper with Koutschan, several Legendre pairs of new lengths $\ell \equiv 0 \pmod{3}$, as well as an algorithm that allows one to determine the full spectrum of values for the $\ell/3$ -rd power spectral density value. The importance of Legendre pairs lies in the fact that they constitute a promising avenue to the Hadamard conjecture.

3.20 Formal Proofs for Cylindrical Algebraic Coverings

Gereon Kremer (Stanford University, US)

License  Creative Commons BY 4.0 International license
© Gereon Kremer

Formally verifiable proofs can be used to boost trust in SMT solvers in cases of unsatisfiability. We show how `cvc5` generates such formal proofs for theory calls to the arithmetic theory. So far, we do not produce verifiable proofs for cylindrical algebraic coverings: we present our current approach and discuss challenges and open questions we face.

3.21 Proving Satisfiability in NTA via Unconstrained Optimization and the Topological Degree Test

Enrico Lipparini (Fondazione Bruno Kessler, IT)

License  Creative Commons BY 4.0 International license
© Enrico Lipparini

Non-linear arithmetic over the reals augmented with transcendental functions (NTA) is well-known to be a tough theory, and proving satisfiability in NTA is especially hard. In this talk, I present a novel procedure to tackle this challenge. Our procedure makes use of

two main ingredients: unconstrained optimisation, to generate a set of candidate solutions, and a result from topology called the topological degree test, to guarantee the existence of a model in a bounded region. We implemented the procedure in a prototype tool called `ugotNL`, proposing both an eager and a lazy approach (the former being integrated within the MathSAT SMT solver). Our experimental evaluation over a wide range of benchmarks shows that both our tools outperform existing methods for satisfiable formulas, significantly advancing the whole state of the art for NTA. At the end, I discuss the potential of our ideas for future works.

3.22 Implementation and Application of Chordality Preserving Top-Down Algorithms for Triangular Decomposition

Chenqi Mou (Beihang University – Beijing, CN)

License © Creative Commons BY 4.0 International license
© Chenqi Mou

Joint work of Chenqi Mou, Wenwen Ju, Mingyu Dong

Recently chordal graphs have been successfully used in the study of elimination methods like triangular decomposition and quantifier elimination. In this talk, I will first briefly review the underlying theories for studying top-down algorithms for triangular decomposition based on chordal graphs. Next I will talk about our implementations of chordality preserving sparse algorithms for triangular decomposition based on the Epsilon package for the computer algebra system Maple. These algorithms are then applied to exploit the variable sparsity of biological dynamical systems in computing their equilibria, and the experimental results will be reported.

3.23 Dynamical System Verification using Abstractions and Non-Linear Arithmetic

Sergio Mover (Ecole Polytechnique – Palaiseau, FR)

License © Creative Commons BY 4.0 International license
© Sergio Mover

A semi-algebraic abstraction is a qualitative abstraction that partitions the state-space of a continuous dynamical system according to the sign of polynomials, obtaining as a result a finite-state, discrete transition system. Standard model checking techniques can then verify safety properties on such discrete system and, if that is the case, conclude safety for the original continuous system.

We can construct such abstraction that for polynomial dynamical systems using Non-Linear Real Arithmetic (NRA). However, the number of discrete states we have to enumerate is, in the worst case, exponential in the number of polynomials. We show how to avoid the up-front computation of the abstraction and, instead, provide an implicit encoding of the abstraction in a discrete transition system expressed symbolically with NRA formulas. We then solve the safety verification problem model checking such transition system.

We conclude evaluating the new technique and discussing some challenges to scale the verification and extend the encoding to hybrid, instead of dynamical, systems.

3.24 Levelwise Construction of a Single Cylindrical Algebraic Cell

Jasper Nalbach (RWTH Aachen University, DE)

License  Creative Commons BY 4.0 International license
 © Jasper Nalbach

Joint work of Jasper Nalbach, Erika Ábrahám, Philippe Specht, Christopher W. Brown, James H. Davenport, Matthew England

Satisfiability modulo theories (SMT) solving is a technique for checking the satisfiability of quantifier-free first-order logic formulas over different theories. We consider the theory of non-linear real arithmetic where the formulae are logical combinations of polynomial constraints. The most commonly used decision procedure is the cylindrical algebraic decomposition (CAD) which has doubly exponential complexity (in the number of real variables). It works through a projection and lifting framework, where the projection tracks the resultants, discriminants and coefficients of all polynomials involved.

A CAD encodes more information than necessary for checking satisfiability. Hence there has been the recent development of sample-based algorithms which reduce the computational effort in CAD by guessing samples and generalizing conflicts by constructing truth-invariant cells around conflicting samples. The most notable example of this is the NLSAT algorithm, an instantiation of the model-constructing satisfiability calculus (MCSAT). Conflict generalization improves on CAD by reducing the number of polynomials to project, and reducing the projection itself based on the sample for the conflict. The original NLSAT only reduces the number of coefficients tracked in projection, but it was followed by the single cell construction which reduces further the number of resultants and discriminants used. This construction involved refining a cell iteratively, polynomial by polynomial, meaning the shape and size of the resulting cell depends on the order in which the polynomials are considered.

Our paper (to be submitted) further develops these ideas by employing a levelwise approach to cell construction, so called as the cell is built level by level according to a variable ordering, rather than incrementally refined according to a polynomial ordering. We still use a reduced number of projection polynomials based on the sample being generalised, but we consider at once all the possibilities for these at a given level allowing for the use of heuristics to select the polynomials used to try and optimise on the shape of the cell

A further contribution is that we have formulated the necessary theory as a proof system, allowing for a decoupling of such heuristic decisions from the main algorithm and its proof of correctness.

Based on a first implementation, we validate experimentally the benefit of this levelwise approach. We compare three basic heuristics and observe that each heuristic has strength on different subsets of the dataset compared to the others; offering clear potential for further exploitation of the new approach.

3.25 Proof theory and computational algebra

Thomas Powell (University of Bath, GB)

License  Creative Commons BY 4.0 International license
 © Thomas Powell

I gave a high-level overview of proof interpretations and their role in applied proof theory, briefly describing their use in both proof mining and formal program synthesis. I then presented some recent work on applying proof theory in abstract algebra and outlined what I

consider to be some interesting discussion points on the potential intersection of proof theory with the Symbolic Computation and Satisfiability Checking community. These include:

1. Are there interesting proof systems suited to reasoning about proofs and programs in computer algebra and satisfiability checking?
2. Can proof theoretic techniques be used as part of a verification strategy for key algorithms in these areas?
3. Are there examples of nonconstructive proofs where proof interpretations can yield useful quantitative information?

3.26 Looking Backward and Forward

Stefan Ratschan (The Czech Academy of Sciences – Prague, CZ)

License  Creative Commons BY 4.0 International license
© Stefan Ratschan

In the talk, I looked back at the influence of some principles from numerical analysis on the development of decision procedures for the theory of real closed fields, discussed how the topological degree can be used as an existence test in this context, and presented some ideas for future work.

3.27 Computational Limits of Using CAD and SMT Solvers in Logical Analysis of Regulatory Networks

AmirHosein Sadeghimanesh (Coventry University, GB)

License  Creative Commons BY 4.0 International license
© AmirHosein Sadeghimanesh

The Multistationarity region of chemical reaction networks can be described as a system of polynomial inequalities on the parameters of the network. However the algebraic algorithms to do so have high computational complexity. Biologists usually break the model to smaller pieces and then describe the behavior of the big model as a function of the behavior of the smaller compartments of the model. This talk tries to start a discussion for creating new algorithms using CAD and SMT-solvers to find the possible boolean formulae expressing multistationarity of a large network according to multistationarity of its smaller compartments if any exists.

3.28 Expansion-Based QBF Solving for QBF

Martina Seidl (Johannes Kepler Universität Linz, AT)

License  Creative Commons BY 4.0 International license
© Martina Seidl

Within the last years, several different solving approaches for quantified Boolean formulas (QBF) have been presented. One particular successful approach is based on the expansion of quantifiers. In this talk, we present a novel algorithm for expansion-based QBF solving.

3.29 Decidable Logics with Arithmetic and Uninterpreted Symbols for SMT

Baptiste Vergain (University of Liège, BE)

License  Creative Commons BY 4.0 International license
© Baptiste Vergain

Joint work of Bernard Boigelot, Pascal Fontaine, Baptiste Vergain

The objective of this *work in progress* is to characterize decidable sublogics of the difference logic with uninterpreted predicates (e.g. UFIDL, UFRDL), i.e. formulas with first-order quantifiers, uninterpreted predicates, and atomic arithmetic expressions of the form $x - y \bowtie c$, where \bowtie ranges over the relations $<$, \leq , $=$, \neq , \geq , $>$, where the intended domain is natural numbers, integers or real numbers.

In particular, we identified the fragment on integers where only unary predicates are allowed, the decidability of which can be established by stating its equivalence with Büchi automata (similarly to S1S), with an effective decision procedure. Our next objective is to adapt this decision procedure to the integers and to the reals.

It is straightforward for integers. The case of real numbers is although much harder. It is not clear to us yet what restrictions are necessary to ensure decidability of the fragment. We discuss the expressive power of the logic over this domain, and introduce an effective representations of the models of a formula based on words indexed by linear orderings. We then investigate how automata theory can lead us towards a decision procedure.

Participants

- Erika Abraham
RWTH Aachen University, DE
- Christopher W. Brown
U.S. Naval Academy –
Annapolis, US
- Alessandro Cimatti
Bruno Kessler Foundation –
Trento, IT
- James H. Davenport
University of Bath, GB
- Tereso del Rio
Coventry University, GB
- Matthew England
Coventry University, GB
- Jürgen Gerhard
Maplesoft – Waterloo, CA
- Alberto Griggio
Bruno Kessler Foundation –
Trento, IT
- Tudor Jebelean
Universität Linz, AT
- Konstantin Korovin
University of Manchester, GB
- Gereon Kremer
Stanford University, US
- Corin Lee
University of Bath, GB
- Sergio Mover
Ecole Polytechnique –
Palaiseau, FR
- Jasper Nalbach
RWTH Aachen University, DE
- Thomas Powell
University of Bath, GB
- Stefan Ratschan
The Czech Academy of Sciences –
Prague, CZ
- Ali K. Uncu
University of Bath, GB
- Baptiste Vergain
University of Liège, BE



Remote Participants

- Haniel Barbosa
Federal University of Minas
Gerais-Belo Horizonte, BR
- Anna Maria Bigatti
University of Genova, IT
- Russell Bradford
University of Bath, GB
- Martin Brain
City – University of London, GB
- Curtis Bright
University of Windsor, CA
- Martin Bromberger
MPI für Informatik –
Saarbrücken, DE
- Changbo Chen
Chinese Academy of Sciences –
Chongqing, CN
- Isabela Dramnesc
West University of
Timisoara, RO
- Bruno Dutertre
Amazon – Cupertino, US
- Madalina Erascu
West University of
Timisoara, RO
- Pascal Fontaine
University of Liège, BE
- Vijay Ganesh
University of Waterloo, CA
- Ahmed Irfan
Amazon – Cupertino, US
- Manuel Kauers
Johannes Kepler Universität
Linz, AT
- Daniela Kaufmann
Johannes Kepler Universität
Linz, AT
- Ilias S. Kotsireas
Wilfrid Laurier University –
Waterloo, CA
- Enrico Lipparini
Bruno Kessler Foundation –
Trento, IT

- Chenqi Mou
Beihang University – Beijing, CN
- Mizuhito Ogawa
JAIST – Ishikawa, JP
- AmirHosein Sadeghimanesh
Coventry University, GB
- Martina Seidl
Johannes Kepler Universität
Linz, AT

Theory of Randomized Optimization Heuristics

Anne Auger^{*1}, Carlos M. Fonseca^{*2}, Tobias Friedrich^{*3},
Johannes Lengler^{*4}, and Armand Gissler^{†5}

- 1 INRIA Saclay – Palaiseau, FR. anne.auger@inria.fr
- 2 University of Coimbra, PT. cmfonsec@dei.uc.pt
- 3 Hasso-Plattner-Institut, Universität Potsdam, DE. friedrich@hpi.de
- 4 ETH Zürich, CH. johannes.lengler@inf.ethz.ch
- 5 Ecole Polytechnique – Palaiseau, FR. armand.gissler@ens-paris-saclay.fr

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 22081 “Theory of Randomized Optimization Heuristics”.

This seminar is part of a biennial seminar series. This year, we focused on connections between classical topics of the community, such as Evolutionary Algorithms and Strategies (EA, ES), Estimation-of-Distribution Algorithms (EDA) and Evolutionary Multi-Objective Optimization (EMO), and related fields like Stochastic Gradient Descent (SGD) and Bayesian Optimization (BO). The mixture proved to be extremely successful. Already the first talk turned into a two hour long, vivid and productive plenary discussion. The seminar was smaller than previous versions (due to corona regulations), but its intensity more than made up for the smaller size.

Seminar February 20–25, 2022 – <http://www.dagstuhl.de/22081>

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, derivative-free optimization, evolutionary and genetic algorithms, randomized search algorithms, stochastic gradient descent, theoretical computer science

Digital Object Identifier 10.4230/DagRep.12.2.87

1 Executive Summary

Anne Auger
Carlos M Fonseca
Tobias Friedrich
Johannes Lengler

License  Creative Commons BY 4.0 International license
© Anne Auger, Carlos M Fonseca, Tobias Friedrich, Johannes Lengler

This seminar is part of a biennial seminar series. This year, we focused on connections between classical topics of the community, such as Evolutionary Algorithms and Evolutionary Strategies (EA, ES), Estimation-of-Distribution Algorithms (EDA) and Evolutionary Multi-Objective Optimization (EMO), and related fields like Stochastic Gradient Descent (SGD) and Bayesian Optimization (BO). The mixture proved to be extremely successful. Already the first talk turned into a two hours long, vivid and productive plenary discussion. Participants

* 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. 12, Issue 2, pp. 87–102

Editors: Anne Auger, Carlos M Fonseca, Tobias Friedrich, Johannes Lengler, and Armand Gissler



Dagstuhl Reports

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

like Peter Richtarik and Sebastian Stich brought valuable new perspectives from the SGD community, and Mickaël Binois contributed the BO perspective. This yielded some new approaches to long-standing open problems, specifically for a convergence proof of the CMA-ES algorithm on quadratic functions.

Another interesting and fruitful aspect of the seminar was a shift of perspective to search spaces that are under-represented in the community. Traditionally, the search spaces are product spaces, either discrete (especially the n -dimensional hypercube), or continuous (d -dimensional Euclidean space). This year we had some intense discussions in plenum and in working groups on other search spaces, triggered especially by Ekhine Irurozki's presentation on permutation spaces.

Naturally, a big part of the seminar was also devoted to classical topics of the community. Highlights included talks by Benjamin Doerr on the first runtime result for the Non-Dominated Sorting Genetic Algorithm (NSGA-II) and by Tobias Glasmachers on Convergence Analysis of the Hessian Estimation Evolution Strategy (HE-ES). The latter is the first convergence proof for a covariance matrix algorithm that does not truncate the condition number of the estimated covariance matrix. Some interesting new topics were also identified in traditional fields, such as whether we can understand better in which situations adaptivity is necessary for efficient optimization by considering k -adaptive query complexity of optimization benchmarks.

Overall, as organizers we were extremely happy with the mix of core community members and researchers from related fields. The connections with the latter were close enough that scientific discussions could (also) happen on technical levels, which is particularly useful since some low-hanging fruits are available from such interchanges. Importantly, the exchange happened between people who would probably not have met each other outside of the Dagstuhl Seminar.

The seminar took place during the peak of the Omicron wave of Covid19, which made planning very difficult. The key step during preparation phase was a survey among the participants a few weeks before the seminar. We asked how likely it was that they could participate in person, and under which circumstances they would prefer which format (in-person or hybrid). The participants signalled us very clearly that they wanted this event to happen, and that they wanted it to happen in person. We want to thank all participants for their support! Other seminars in the week before and after ours had to be cancelled altogether, and this might also have happened to our seminar if not for the determination of our participants.

The seminar was smaller than previous versions, due to corona regulations. Moreover, some participants had to cancel at the last moment because they were corona-positive, or because they had no reliable child care. Especially the latter point can be frustrating, and we hope that Dagstuhl will be able to resume their support for on-site child care in the future. On the positive side, the intensity of the seminar more than made up for the smaller size, and might even have been due to the smaller number of participants.

Finally, we want to thank Dagstuhl for their great support, both financially and to their great staff. We could always feel that it was their top priority to help us, and we are greatly indebted for the support!

The organizers,

Anne Auger, Carlos M Fonseca, Tobias Friedrich, Johannes Lengler

2 Table of Contents

Executive Summary

Anne Auger, Carlos M Fonseca, Tobias Friedrich, Johannes Lengler 87

Overview of Talks

Selection in non-elitist populations: overview and open problems <i>Duc-Cuong Dang</i>	91
A First Mathematical Runtime Analysis of the Non-Dominated Sorting Genetic Algorithm II (NSGA-II) <i>Benjamin Doerr</i>	91
Some Theoretical Thoughts on Permutation-based EAs <i>Benjamin Doerr</i>	92
k -Adaptive Black-Box Optimization <i>Carola Doerr</i>	92
Convergence Analysis of the Hessian Estimation Evolution Strategy <i>Tobias Glasmachers</i>	93
Evolution Strategies Reliably Overcome Saddle Points <i>Tobias Glasmachers</i>	93
Introductory Talk on the Theory of Continuous Evolutionary Algorithms <i>Tobias Glasmachers</i>	93
Theoretical Aspects of Set-Quality Indicators for Multiobjective Optimization <i>Andreia P. Guerreiro</i>	94
Black-Box Permutation Problems and weighted medians <i>Ekhine Irurozki</i>	94
Theory of Discrete Randomized Optimization Heuristics – The What, Why, and How <i>Martin S. Krejca</i>	95
Failure on Easy Problems <i>Johannes Lengler, Benjamin Doerr, Carola Doerr, and Dirk Sudholt</i>	95
Population Diversity Makes Lexicase Selection Fast <i>Johannes Lengler</i>	96
Gray-box operator for Vertex Cover <i>Xiaoyue Li</i>	96
Randomized Smoothing for Non-Convex Optimization <i>Sebastian U. Stich</i>	97
Analyzing the Cost of Randomness in Evolutionary Algorithms <i>Dirk Sudholt and Carlo Kneißl</i>	97
On the dynamics of the DE algorithm <i>Ricardo Takahashi</i>	98
Mathematical models for Dominance Move: Comparisons and complexity analysis <i>Elizabeth Wanner</i>	98

Working groups

<i>k</i> -adaptive black-box complexity <i>Carola Doerr and Johannes Lengler</i>	98
Theory-Friendly Practical Modelling of Combinatorial Optimisation Problems for ROHs <i>Carlos M. Fonseca</i>	99
Adaptation of proof techniques based on the natural SGD interpretation of CMA-ES <i>Tobias Glasmachers</i>	100
Lyapunov potential for the linear convergence of CMA-ES <i>Tobias Glasmachers</i>	100
Comma and plus selection strategies with noise <i>Johannes Lengler</i>	100

Open problems

A problem where CMA-ES performs poorly <i>Nikolaus Hansen</i>	101
--	-----

Participants	102
-------------------------------	-----

3 Overview of Talks

3.1 Selection in non-elitist populations: overview and open problems

Duc-Cuong Dang (University of Southampton, GB)

License © Creative Commons BY 4.0 International license
© Duc-Cuong Dang

Joint work of Duc-Cuong Dang, Anton V. Eremeev, Per Kristian Lehre

This talk summarises what we know about the selection mechanisms for non-elitist populations from a theory perspective and open problems. Particularly, we show how selection should be tuned to find an optimum, and which characteristics of the selection one should look for to address multiple optima efficiently. These results allow us to identify problem classes where non-elitist algorithms with the right selection and a proper setting to excel, compared to elitist algorithms or even when truncation selection is used. Tools used to prove these results, these limitations and open problems are discussed.

This talk is based on joint works with Anton V. Eremeev and Per Kristian Lehre [1, 2, 3, 4], and on the fruitful discussions with Pietro S. Oliveto and Tiago Paixao.

References

- 1 D. Corus and D. Dang and A.V. Eremeev and P.K. Lehre *Level-based analysis of genetic algorithms and other search processes*. IEEE Trans. Evol. Comput., 707–719, 2018.
- 2 D. Dang and A.V. Eremeev and P.K. Lehre *Runtime analysis of fitness-proportionate selection on linear*. <http://arxiv.org/abs/1908.08686>, 2019.
- 3 D. Dang and A.V. Eremeev and P.K. Lehre *Non-elitist evolutionary algorithms excel in fitness landscapes with sparse deceptive regions and dense valleys*. In GECCO '21: Genetic and Evolutionary Computation Conference, Lille, France, 1133–1141, July 2021.
- 4 D. Dang and A.V. Eremeev and P.K. Lehre *Escaping local optima with non-elitist evolutionary algorithms*. Thirty-Fifth AAAI Conference on Artificial Intelligence, 12275–12283, 2021.

3.2 A First Mathematical Runtime Analysis of the Non-Dominated Sorting Genetic Algorithm II (NSGA-II)

Benjamin Doerr (Ecole Polytechnique – Palaiseau, FR)

License © Creative Commons BY 4.0 International license
© Benjamin Doerr

Joint work of Weijie Zheng, Yufei Liu, Benjamin Doerr

Main reference Weijie Zheng, Yufei Liu, Benjamin Doerr: “A First Mathematical Runtime Analysis of the Non-Dominated Sorting Genetic Algorithm II (NSGA-II)”, CoRR, Vol. abs/2112.08581, 2021.

URL <https://arxiv.org/abs/2112.08581>

In this talk, I want to discuss a recent joint work with Weijie Zheng (SUSTECH) and Yufei Liu (Polytechnique). The non-dominated sorting genetic algorithm II (NSGA-II) is the most intensively used multi-objective evolutionary algorithm (MOEA) in real-world applications. However, in contrast to several simple MOEAs analyzed also via mathematical means, no such study exists for the NSGA-II so far. In this work, we show that mathematical runtime analyses are feasible also for the NSGA-II. As particular results, we prove that with a population size larger than the Pareto front size by a constant factor, the NSGA-II with two classic mutation operators and three different ways to select the parents satisfies the same asymptotic runtime guarantees as the SEMO and GSEMO algorithms on the basic

OneMinMax and LOTZ benchmark functions. However, if the population size is only equal to the size of the Pareto front, then the NSGA-II cannot efficiently compute the full Pareto front (for an exponential number of iterations, the population will always miss a constant fraction of the Pareto front). Our experiments confirm the above findings.

3.3 Some Theoretical Thoughts on Permutation-based EAs

Benjamin Doerr (Ecole Polytechnique – Palaiseau, FR)

License  Creative Commons BY 4.0 International license
 © Benjamin Doerr

Joint work of Benjamin Doerr, Yassine Ghannane, Marouane Ibn Brahim

While the theoretical analysis of evolutionary algorithms (EAs) has made significant progress for pseudo-Boolean optimization problems in the last 25 years, only sporadic theoretical results exist on how EAs solve permutation-based problems.

To overcome the lack of permutation-based benchmark problems, we propose a general way to transfer the classic pseudo-Boolean benchmarks into benchmarks defined on sets of permutations. We then conduct a rigorous runtime analysis of the permutation-based $(1 + 1)$ EA proposed by Scharnow, Tinnefeld, and Wegener (2004) on the analogues of the LEADINGONES and JUMP benchmarks. The latter shows that, different from bit-strings, it is not only the Hamming distance that determines how difficult it is to mutate a permutation σ into another one τ , but also the precise cycle structure of $\sigma\tau^{-1}$. For this reason, we also regard the more symmetric scramble mutation operator. We observe that it not only leads to simpler proofs, but also reduces the runtime on jump functions with odd jump size by a factor of $\Theta(n)$. Finally, we show that heavy-tailed versions of both operators, as in the bit-string case, lead to speed-ups of order $m^{\Theta(m)}$ on jump functions with jump size m .

3.4 k -Adaptive Black-Box Optimization

Carola Doerr (Sorbonne University – Paris, FR)

License  Creative Commons BY 4.0 International license
 © Carola Doerr

Many black-box optimization techniques have a high degree of adaptiveness. But there are problems for which adaptive sampling has only negligible advantages over non-adaptive sampling, e.g., the famous 2-color Mastermind problem studied by Erdős and Renyi (1963).

In this talk I propose a black-box complexity model that allows us to study the minimal number of queries that are needed to optimize a given problem f , in dependency of the number of iterations k that the algorithms are allowed to perform.

3.5 Convergence Analysis of the Hessian Estimation Evolution Strategy

Tobias Glasmachers (Ruhr-Universität Bochum, DE)

License © Creative Commons BY 4.0 International license
© Tobias Glasmachers

Joint work of Tobias Glasmachers, Oswin Krause

Main reference Tobias Glasmachers, Oswin Krause: “Convergence Analysis of the Hessian Estimation Evolution Strategy”, *Evol. Comput.*, Vol. 30(1), pp. 27–50, 2022.

URL https://doi.org/10.1162/evco_a_00295

I will sketch the convergence proof of a minimal elitist variant of the recently proposed Hessian Estimation Evolution Strategy (HE-ES). The main difference as compared with CMA-ES is that the covariance matrix update yields monotonic convergence of the covariance matrix to the inverse Hessian. This strong stability property allows to prove that the algorithm converges at a linear rate to the minimum of a convex quadratic objective function, where the rate is independent of the problem instance. The same holds for CMA-ES, but we are lacking a proof (since 20 years). The proof works in two steps, both of which employ drift arguments. The first step is to prove convergence of the covariance matrix, which works independent of the evolution of mean and step size. The second step is to reduce the convergence proof to recent powerful results for the (1+1)-ES without CMA.

3.6 Evolution Strategies Reliably Overcome Saddle Points

Tobias Glasmachers (Ruhr-Universität Bochum, DE)

License © Creative Commons BY 4.0 International license
© Tobias Glasmachers

I will present the (to my knowledge) first result on the behavior of an ES facing saddle points. The (1+1)-ES overcomes even rather difficult saddle points, where it could be subjected to convergence prematurely because the success rate is smaller than $1/5$, and in fact arbitrarily close to zero.

3.7 Introductory Talk on the Theory of Continuous Evolutionary Algorithms

Tobias Glasmachers (Ruhr-Universität Bochum, DE)

License © Creative Commons BY 4.0 International license
© Tobias Glasmachers

I will introduce basic problems (in particular convex quadratic functions) and algorithms (the classic (1+1)-ES and a variant with covariance matrix adaptation (CMA)). I will describe the qualitative behavior of evolution strategies with and without CMA on problems with good and bad conditioning. Some aspects of the behavior are well described by theoretical analysis, while others are open problems. I’ll then give a gist of existing analysis methodologies: Markov chains (and a recent way of proving stability), drift, and the IGO framework.

3.8 Theoretical Aspects of Set-Quality Indicators for Multiobjective Optimization

Andreia P. Guerreiro (INESC-ID – Lisboa, PT)

License  Creative Commons BY 4.0 International license
© Andreia P. 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.

3.9 Black-Box Permutation Problems and weighted medians

Ekhine Irurozki (Telecom Paris, FR)

License  Creative Commons BY 4.0 International license
© Ekhine Irurozki

Main reference Ekhine Irurozki, Manuel López-Ibáñez: “Unbalanced mallows models for optimizing expensive black-box permutation problems”, in Proc. of the GECCO '21: Genetic and Evolutionary Computation Conference, Lille, France, July 10-14, 2021, pp. 225–233, ACM, 2021.

URL <https://doi.org/10.1145/3449639.3459366>

From the paper: Unbalanced Mallows Models for Optimizing Expensive Black-Box Permutation Problems

Expensive black-box combinatorial optimization problems arise in practice when the objective function is evaluated by means of a simulator or a real-world experiment. Since each fitness evaluation is expensive in terms of time or resources, the number of possible evaluations is typically several orders of magnitude smaller than in non-expensive problems. Classical optimization methods are not useful in this scenario. In this talk, we propose and analyze UMM, an estimation-of-distribution (EDA) algorithm based on a Mallows probabilistic model and unbalanced rank aggregation (uBorda). UMM is based on a weighted median for permutations. The core of it is uBorda. Experimental results on black-box versions of LOP and PFSP show that UMM outperforms the solutions obtained by CEGO, a Bayesian optimization algorithm for combinatorial optimization. Nevertheless, a slight modification to CEGO, based on the different interpretations for rankings and orderings, significantly improves its performance, thus producing solutions that are slightly better than those of UMM and dramatically better than the original version. Another benefit of UMM is that its computational complexity increases linearly with both the number of function evaluations and the permutation size, which results in computation times an order of magnitude shorter than CEGO, making it specially useful when both computation time and number of evaluations are limited.

3.10 Theory of Discrete Randomized Optimization Heuristics – The What, Why, and How

Martin S. Krejca (Sorbonne University – Paris, FR)

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

Randomized optimization heuristics (ROHs) are algorithms applied to optimization problems where the objective function is only indirectly accessible, that is, it can only be accessed by evaluating solution candidates. Guided by the quality of such candidates, ROHs aim to iteratively generate solutions of better quality. This raises natural questions such as how quickly an ROH improves its solutions, whether it is capable of finding optimal solutions, or if there are any approaches that generate better solutions more quickly. Theoretical analyses aim to answer these questions and more.

In this talk, we provide an introduction to theoretical analyses on ROHs in the *discrete* domain. We introduce the setting that ROHs are applied in, the so-called *black-box* setting, and we discuss how theoretical results aim to answer questions about the real-world application of ROHs. To this end, we introduce common theory benchmark functions, ROHs, as well as the mathematical tools used for their analyses. We conclude by deriving a standard run time result, illustrating how problems in this domain are typically approached.

3.11 Failure on Easy Problems

Johannes Lengler (ETH Zürich, CH), Benjamin Doerr (Ecole Polytechnique – Palaiseau, FR), Carola Doerr (Sorbonne University – Paris, FR), and Dirk Sudholt (Universität Passau, DE)

License  Creative Commons BY 4.0 International license
© Johannes Lengler, Benjamin Doerr, Carola Doerr, and Dirk Sudholt

I will discuss several seemingly easy situations in which evolutionary and genetic algorithms can fail. Some of them are rather surprising, since the hardest regions are not always close to the optimum. In particular, I will mention the detrimental effects of large mutation rates or of large population size for optimizing monotone functions, introduce a simple dynamic setting, and discuss how the self-adapting $(1, \lambda)$ -EA can fail on OneMax.

References

- 1 Benjamin Doerr, Thomas Jansen, Dirk Sudholt, Carola Winzen and Christine Zarges. *Mutation rate matters even when optimizing monotonic functions*. Evolutionary computation, 21(1), 2013, 1-27.
- 2 Johannes Lengler. *A general dichotomy of evolutionary algorithms on monotone functions*. IEEE Transactions on Evolutionary Computation, 24(6), 995–1009, 2019, IEEE
- 3 Johannes Lengler and Jonas Meier. *Large population sizes and crossover help in dynamic environments*. In: International Conference on Parallel Problem Solving from Nature. Springer, Cham, 2020. S. 610-622.
- 4 Mario Alejandro Hevia Fajardo and Dirk Sudholt. *Self-adjusting population sizes for non-elitist evolutionary algorithms: why success rates matter*. In: Proceedings of the Genetic and Evolutionary Computation Conference. 2021. S. 1151-1159.
- 5 Johannes Lengler and Xun Zou. *Exponential slowdown for larger populations: The $(\mu+1)$ -EA on monotone functions*. Theoretical Computer Science 875 (2021): 28-51.

- 6 Johannes Lengler and Simone Riedi. *Runtime Analysis of the $(\mu + 1)$ -EA on the Dynamic BinVal Function*. In: European Conference on Evolutionary Computation in Combinatorial Optimization (Part of EvoStar). Springer, Cham, 2021. S. 84-99.

3.12 Population Diversity Makes Lexicase Selection Fast

Johannes Lengler (ETH Zürich, CH)

License  Creative Commons BY 4.0 International license
© Johannes Lengler

Joint work of William de Casa, Thomas Helmuth, Johannes Lengler

In Genetic Programming, it is customary to have population sizes in the order of $\mu = 100 - 1000$. For the next generation, $\lambda = 2\mu$ parents are independently selected. In order to select a parent, lexicase selection first removes duplicates in performance space, i.e., individuals which show the same performance on all test cases. Then it picks one test case at random, and removes all candidates which fail on this test case (unless all candidates fail, in which case the test is skipped). This is iterated until only one candidate remains. My two co-authors report that lexicase selection is regarded critically in their community due to its very bad worst-case runtime, but that the runtime in practice is much faster. We investigated why the runtime is fast and found a measure for population diversity such that i) empirically population diversity is large, and ii) a large population diversity guarantees theoretically fast runtimes.

3.13 Gray-box operator for Vertex Cover

Xiaoyue Li (Hasso-Plattner-Institut, Universität Potsdam, DE)

License  Creative Commons BY 4.0 International license
© Xiaoyue Li

Joint work of Samuel Baguley, Tobias Friedrich, Timo Kötzing, Xiaoyue Li, Marcus Pappik, Ziena Zeif
Main reference Samuel Baguley, Tobias Friedrich, Timo Kötzing, Xiaoyue Li, Marcus Pappik, Ziena Zeif: “Analysis of a gray-box operator for vertex cover”, in Proc. of the GECCO ’22: Genetic and Evolutionary Computation Conference, Boston, Massachusetts, USA, July 9 – 13, 2022, pp. 1363-1371, ACM, 2022.
URL <https://doi.org/10.1145/3512290.3528848>

In this flash talk, a gray-box operator tailored for combinatorial optimization problems is presented. The operator is called balanced flip EA, which the mutation operator has been introduced to the specific problem so that the algorithm is extended for better behavior. By applying runtime analysis of both the $(1+1)$ EA and the balanced EA, a tighter upper bound is provided for balanced flip EA as $\mathcal{O}(\sqrt{\lambda})$ while the $(1+1)$ EA is as $\mathcal{O}(\sqrt{\lambda})$. Other than the introduction of analysis strategy, the experimental evidence for long bad path to support the partial proof of the lower bound as well as the probability bound for complete bipartite graph has also been discussed. As a result, the flash talk presented the key point that the benefits of gray-box operator should be considered. Especially when it comes to solve the specific problem like combinatorial optimization.

3.14 Randomized Smoothing for Non-Convex Optimization

Sebastian U. Stich (CISPA – Saarbrücken, DE)

License © Creative Commons BY 4.0 International license
© Sebastian U. Stich

Joint work of Ahmad Ajalloeian, Harsh Vardhan, Sebastian Urban Stich

We identify a class of nonconvex functions for which we can show that perturbed gradient descent converges to a global minimum, in contrast to gradient descent without noise that can get stuck in local minima far from a global solution.

We give a brief overview of the used techniques, such as convergence analysis of (stochastic) gradient methods, biased gradients, and randomized smoothing.

Based on joint work with A. Ajalloeian [1] and H. Vardhan [2].

References

- 1 A. Ajalloeian and S.U. Stich *On the Convergence of SGD with Biased Gradients*. presented at ICML 2020 Workshop on *Beyond First Order Methods in ML Systems*, arXiv preprint arXiv:2008.00051, 2021.
- 2 H. Vardhan and S.U. Stich *Tackling benign nonconvexity with smoothing and stochastic gradients*. presented at NeurIPS 2021 Workshop on *Optimization for Machine Learning*, arXiv preprint arXiv:2202.0905, 2022.

3.15 Analyzing the Cost of Randomness in Evolutionary Algorithms

Dirk Sudholt (Universität Passau, DE) and Carlo Kneißl

License © Creative Commons BY 4.0 International license
© Dirk Sudholt and Carlo Kneißl

Evolutionary algorithms make countless random decisions during selection, mutation and crossover operations. These random decisions require a steady stream of random numbers, however generating good quality randomness is non-trivial.

We consider the expected number of random bits used throughout a run of an evolutionary algorithm and refer to this as the cost of randomness. We give general bounds on the cost of randomness for mutation-based evolutionary algorithms using 1-bit flips or standard bit mutations using either a naive or a common, more efficient implementation that uses $\Theta(\log n)$ random bits per mutation. Uniform crossover is a potentially wasteful operator as the number of random bits used equals the Hamming distance of the two parents, which can be up to n . However, we show for a $(2+1)$ GA that optimizes OneMax in expected $\Theta(n \log n)$ evaluations that the total cost of randomness during all crossover operations on OneMax is only $\Theta(n)$.

We hope to show that the cost of randomness may be useful as an additional performance measure, to give new insights into search dynamics and to aid in the design of operators that use randomness more carefully.

3.16 On the dynamics of the DE algorithm

Ricardo Takahashi (Federal University of Minas Gerais-Belo Horizonte, BR)

License  Creative Commons BY 4.0 International license
© Ricardo Takahashi

This talk presents an ongoing work that aims to develop an analytical study of the Differential Evolution (DE) algorithm behavior. Analytical formulae for the probability of enhancement of and individual in populations of the DE/rand/1/bin and DE/rand/1/exp algorithm versions are developed for the sphere objective function. It is shown that those formulae can be adapted for the study of the algorithm behavior in the optimization of quadratic functions with different relations between the minimal and maximal eigenvalues. In the case of large differences of eigenvalue magnitudes, it is shown that the convergence occurs in different scales, and the formulae approximately hold for the corresponding scale dimension. The known effect of performance degradation as the problem dimension increases is partly explained by the decrease in the probability of enhancement of individuals, as indicated in the formulae. Experimental results show that DE/best algorithm versions, as expected, are able to solve problems of higher dimensions. Further research will be performed in order to examine the convergence dependency with the algorithm parameters.

3.17 Mathematical models for Dominance Move: Comparisons and complexity analysis

Elizabeth Wanner (CEFET – Belo Horizonte, BR)

License  Creative Commons BY 4.0 International license
© Elizabeth Wanner

Dominance move (DoM), a binary quality indicator, can be used in multi-objective and many-objective optimization to compare two solution sets. DoM is very intuitive but hard to calculate due to its combinatorial nature. Different mathematical models are presented and analyzed. A computationally fast approximate approach is also discussed. Computational results are promising and an upper bound analysis for the approximation ratio would be useful.

4 Working groups

4.1 k -adaptive black-box complexity

Carola Doerr (Sorbonne University – Paris, FR) and Johannes Lengler (ETH Zürich, CH)

License  Creative Commons BY 4.0 International license
© Carola Doerr and Johannes Lengler

In k -adaptive black-box complexity, algorithms are only allowed to perform k iterations of queries through which the optimal solution needs to be learned. In each iteration, the algorithms are allowed to perform an arbitrary number of queries. We are interested in the minimal total number of queries that a k -adaptive black-box optimization algorithm needs

to perform in order to find an optimal solution. The k -adaptive black-box complexity models thus interpolate between non-adaptive query complexity ($k = 1$) and fully adaptive query complexity (no restriction on k).

We discussed the k -adaptive black box complexity (BBC) for several benchmark problems, starting with OneMax and Mastermind. For OneMax, the query complexity does not substantially increase (only by a constant factor) even if we restrict to $k = 1$. For general Mastermind, this is unclear and poses an interesting research question. We took some first steps in discussing the $k = 2$ case with n colors and positions.

For LeadingOnes, the situation is almost opposite than for OneMax: if $k = n/\alpha$, then we can find an algorithm with $2^\alpha n/\alpha$ queries, and we believe that the query complexity is exponentially large in α , perhaps even $\Omega(2^\alpha n/\alpha)$. We spent some time discussing the differences to Permutation-LeadingOnes: it is at least as hard as LeadingOnes, so the lower bounds still apply. Analyzing previous work of Benjamin Doerr and Carola Doerr [1], we have an algorithm that is n -adaptive and has complexity $O(n \log n / \log \log n)$. Perhaps this could be improved to $O(n \log \log n)$, but the general picture is that adaptivity needs to be very high to avoid an exponentially high price in terms of complexity.

We finally discussed the situation for HiddenSubset, which seems to be an example for a problem that requires an intermediate level of adaptivity. Known algorithms are $(\log n)$ -adaptive, with optimal runtime $O(n \log n)$. It might be interesting to study smaller k , and we conjecture that this would increase the complexity substantially.

References

- 1 Benjamin Doerr and Carola Winzen *Black-Box Complexity: Breaking the $O(n \log n)$ Barrier of LeadingOnes*. presented at Artificial Evolution – 10th International Conference, Evolution Artificielle, EA 2011, Angers, France, October 24-26, 2011, Revised Selected Papers, 205–216, https://doi.org/10.1007/978-3-642-35533-2_18, 2011.
- 2 Jin-Kao Hao and Pierrick Legrand and Pierre Collet and Nicolas Monmarché and Evelyne Lutton and Marc Schoenauer, *Artificial Evolution – 10th International Conference, Evolution Artificielle, EA 2011, Angers, France, October 24-26, 2011, Revised Selected Papers, Lecture Notes in Computer Science, 7401, 2012.*

4.2 Theory-Friendly Practical Modelling of Combinatorial Optimisation Problems for ROHs

Carlos M. Fonseca (University of Coimbra, PT)

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

Problem modelling is often overlooked as the crucial step preceding the practical application of randomised optimisation heuristics (ROHs). In order to make ROHs more accessible in the real world, there is a need to specify how optimisation algorithms interface with the problem instances of interest. In addition, guidance should be provided to practitioners on how problem-specific information can be exposed to the algorithms through such an interface. In other words, such an interface definition should support an associated problem modelling methodology.

From a different perspective, such an interface specification must also support the development of practical algorithms that can be applied directly to any given problem implementing that interface. A brief presentation of two Application Programming Interfaces

(APIs) for combinatorial optimisation under development, and of the design principles behind them, was the starting point for a discussion on the potential of such an approach and how theory-friendly it might be.

4.3 Adaptation of proof techniques based on the natural SGD interpretation of CMA-ES

Tobias Glasmachers (Ruhr-Universität Bochum, DE)

License  Creative Commons BY 4.0 International license
© Tobias Glasmachers

As a minimal example, we went through IGO framework and NES algorithm, using 1D example $f(x) = x^2$ and Gaussians. We discussed the quantile rescaling technique in detail, and whether it may or may not be understood as a static fitness transformation. We also discussed the role of rescaling the gradient with the Fisher information matrix for achieving linear convergence. We identified potential connections to variance-reduced gradient descent techniques, as well as to the sampling literature. The general impression was that the natural SGD setting with quantile rescaling is too far from the standard analysis of SGD to allow for a straightforward transfer of proof techniques.

4.4 Lyapunov potential for the linear convergence of CMA-ES

Tobias Glasmachers (Ruhr-Universität Bochum, DE)

License  Creative Commons BY 4.0 International license
© Tobias Glasmachers

We started with a “naive” potential, which is a straightforward extension of the (1+1)-ES potential. We then focused on Armand’s problem and understood that it is a model of the initial phase of optimizing the discus problem. We tried to fix the $\log(\|m\|)$ term to achieve additive drift in this situation, without much success. Therefore, we shifted the discussion to how a multi-stage drift for that situation may work.

4.5 Comma and plus selection strategies with noise

Johannes Lengler (ETH Zürich, CH)

License  Creative Commons BY 4.0 International license
© Johannes Lengler

Main reference Benjamin Doerr: “Does Comma Selection Help to Cope with Local Optima?”, *Algorithmica*, Vol. 84(6), pp. 1659–1693, 2022.

URL <https://doi.org/10.1007/s00453-021-00896-7>

We started gathering examples of problems on which comma selection performs differently to plus selection, focusing on theoretically proven examples. Although the topic was originally set to deal with noise, we considered deterministic problems for the most part.

It is known that comma selection in a (μ, λ) EA does not perform better than plus selection on JUMP functions, and this holds for arbitrary population sizes μ and λ . On the CLIFF function, comma strategies perform well, enabling a $(1, \lambda)$ EA to jump down the cliff if

all offspring are down the cliff and the algorithm moves towards the global optimum without jumping back up the cliff. For optimal fixed λ , the best known expected runtime is essentially $n^{3.97\dots}$ (up to sub-polynomial factors). Self-adjusting λ can reduce this to $O(n \log n)$ if a reset mechanism is used that resets λ to 1 if it exceeds a given threshold. The effect is similar to using hyperheuristics or ageing where occasionally non-elitist steps are accepted. The structure of the function is quite benign as the gradient past the cliff points towards the global optimum.

It is also known that for large values of λ , a comma strategy will behave like a plus strategy because there is a high probability of cloning the parent. If λ is too small, $\lambda \leq (1 - \varepsilon) \log_{\frac{e}{e-1}} n$, optimising any fitness function with a unique optimum takes exponential time. So there is a narrow region for values of λ where one can see an advantage of comma selection over plus selection.

We also discussed elitist versus non-elitist algorithms more generally (e.g. tournament selection, self-adaptation and island models). For all algorithms discussed, the crucial issue came down to balancing exploration and exploitation: being able to escape from local optima while also being able to climb hills.

We came to the conclusion that there is a gap between theory and practice as we're lacking convincing examples (apart from CLIFF) where comma selection provably helps, whereas in practice comma strategies seem to be quite popular to escape from local optima. Part of this gap might be caused by theory traditionally aiming to find the exact global optimum, whereas in practice one is usually content with a good approximation of the optimal fitness. Using a fixed-target perspective for less ambitious targets might give a different picture as much smaller values of λ might be sufficient.

5 Open problems

5.1 A problem where CMA-ES performs poorly

Nikolaus Hansen (INRIA Saclay – Palaiseau, FR)

License © Creative Commons BY 4.0 International license
© Nikolaus Hansen

We investigate a 6-dimensional curve fitting problem, where CMA-ES takes a long time to approach the global optimum. The problem is simple and visually intuitive. Parameters move very slowly in the same direction towards the optimum for a long time, in which case we would expect the covariance matrix to speed up the movements. However, the sample distribution appears to be stable with a condition number of about 10^6 . The leading hypothesis for the reason is that the problem has a very narrow but slightly bent ridge. The bend prevents a faster approach to the optimum.

Participants

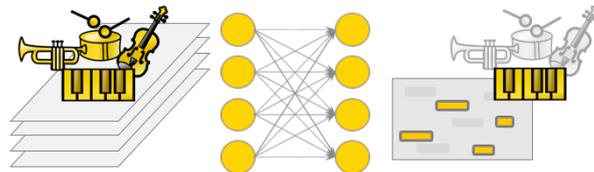
- Mickaël Binois
INRIA – Sophia Antipolis, FR
- Alexandre Chotard
Calais University, FR
- Duc-Cuong Dang
University of Southampton, GB
- Benjamin Doerr
Ecole Polytechnique –
Palaiseau, FR
- Carola Doerr
Sorbonne University – Paris, FR
- Carlos M. Fonseca
University of Coimbra, PT
- Armand Gissler
Ecole Polytechnique –
Palaiseau, FR
- Tobias Glasmachers
Ruhr-Universität Bochum, DE
- Andreia P. Guerreiro
INESC-ID – Lisboa, PT
- Nikolaus Hansen
INRIA Saclay –
Palaiseau, FR
- Ekhnine Irurozki
Telecom Paris, FR
- Martin S. Krejca
Sorbonne University – Paris, FR
- Johannes Lengler
ETH Zürich, CH
- Xiaoyue Li
Hasso-Plattner-Institut,
Universität Potsdam, DE
- Peter Richtarik
KAUST – Thuwal, SA
- Günter Rudolph
TU Dortmund, DE
- Sebastian U. Stich
CISPA – Saarbrücken, DE
- Dirk Sudholt
Universität Passau, DE
- Andrew M. Sutton
University of Minnesota –
Duluth, US
- Ricardo Takahashi
Federal University of Minas
Gerais-Belo Horizonte, BR
- Elizabeth Wanner
CEFET – Belo Horizonte, BR



Deep Learning and Knowledge Integration for Music Audio Analysis

Meinard Müller^{*1}, Rachel Bittner^{*2}, Juhan Nam^{*3},
Michael Krause^{†4}, and Yigitcan Özer^{†5}

- 1 Friedrich-Alexander-Universität Erlangen-Nürnberg, DE.
meinard.mueller@audiolabs-erlangen.de
- 2 Spotify – Paris, FR. rachelbittner@spotify.com
- 3 KAIST – Daejeon, KR. juhan.nam@kaist.ac.kr
- 4 Friedrich-Alexander-Universität Erlangen-Nürnberg, DE.
michael.krause@audiolabs-erlangen.de
- 5 Friedrich-Alexander-Universität Erlangen-Nürnberg, DE.
yigitcan.oezer@audiolabs-erlangen.de



Abstract

Given the increasing amount of digital music, the development of computational tools that allow users to find, organize, analyze, and interact with music has become central to the research field known as Music Information Retrieval (MIR). As in general multimedia processing, many of the recent advances in MIR have been driven by techniques based on deep learning (DL). There is a growing trend to relax problem-specific modeling constraints from MIR systems and instead apply relatively generic DL-based approaches that rely on large quantities of data. In the Dagstuhl Seminar 22082, we critically examined this trend, discussing the strengths and weaknesses of these approaches using music as a challenging application domain. We mainly focused on music analysis tasks applied to audio representations (rather than symbolic music representations) to give the seminar cohesion. In this context, we systematically explored how musical knowledge can be integrated into or relaxed from computational pipelines. We then discussed how this choice could affect the explainability of models or the vulnerability to data biases and confounding factors. Furthermore, besides explainability and generalization, we also addressed efficiency, ethical and educational aspects considering traditional model-based and recent data-driven methods. In this report, we give an overview of the various contributions and results of the seminar. We start with an executive summary describing the main topics, goals, and group activities. Then, we give an overview of the participants' stimulus talks and subsequent discussions (listed alphabetically by the main contributor's last name) and summarize further activities, including group discussions and music sessions.

Seminar February 20–25, 2022 – <http://www.dagstuhl.de/22082>

2012 ACM Subject Classification Information systems → Music retrieval; Applied computing → Sound and music computing

Keywords and phrases Audio signal processing, deep learning, knowledge representation, music information retrieval, user interaction and interfaces

Digital Object Identifier 10.4230/DagRep.12.2.103

* Editor / Organizer

† Editorial Assistant / Collector



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

Deep Learning and Knowledge Integration for Music Audio Analysis, *Dagstuhl Reports*, Vol. 12, Issue 2, pp. 103–133

Editors: Rachel Bittner, Meinard Müller, Juhan Nam, Michael Krause, and Yigitcan Özer



DAGSTUHL
REPORTS Dagstuhl Reports

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

1 Executive Summary

Meinard Müller (Friedrich-Alexander-Universität Erlangen-Nürnberg, DE)

Rachel Bittner (Spotify – Paris, FR)

Juhan Nam (KAIST – Daejeon, KR)

License  Creative Commons BY 4.0 International license
© Meinard Müller, Rachel Bittner, and Juhan Nam

This executive summary gives an overview of our discussions on the integration of musical knowledge in deep learning approaches while summarizing the main topics covered in this seminar. We also describe the seminar’s group composition, the overall organization, and the seminar’s activities. Finally, we reflect on the most important aspects of this seminar and conclude with future implications and acknowledgments.

Overview

Music is a ubiquitous and vital part of our lives. Thanks to the proliferation of digital music services, we have access to music nearly anytime and anywhere, and we interact with music in a variety of ways, both as listeners and active participants. As a result, music has become one of the most popular categories of multimedia content. In general terms, music processing research aims to contribute concepts, models, and algorithms that extend our capabilities of accessing, analyzing, understanding, and creating music. In particular, the development of computational tools that allow users to find, organize, analyze, generate, and interact with music has become central to the research field known as Music Information Retrieval (MIR). Given the complexity and diversity of music, research has to account for various aspects such as the genre, instrumentation, musical form, melodic and harmonic properties, dynamics, tempo, rhythm, timbre, and so on.

As in general multimedia processing, many of the recent advances in MIR have been driven by techniques based on deep learning (DL). For example, DL-based techniques have led to significant improvements for numerous MIR tasks including music source separation, music transcription, chord recognition, melody extraction, beat tracking, tempo estimation, and lyrics alignment. In particular, major improvements could be achieved for specific music scenarios where sufficient training data is available. A particular strength of DL-based approaches is their capability to extract complex features directly from raw audio data, which can then be used for making predictions based on hidden structures and relations. Furthermore, powerful software packages allow for easily designing, implementing, and experimenting with machine learning models based on deep neural networks (DNNs).

However, DL-based approaches also come at a cost, being a data-hungry and computing-intensive technology. Furthermore, the design of suitable network architectures (including the adaption of hyper-parameters and optimization strategies) can be cumbersome and time-consuming – a process that is commonly seen more as an art rather than a science. Finally, the behavior of DL-based systems is often hard to understand; the trained models may capture information that is not directly related to the core problem. These general properties of DL-based approaches can also be observed when analyzing and processing music, which spans an enormous range of forms and styles – not to speak of the many ways music may be generated and represented. While one aims in music analysis and classification problems at capturing musically relevant aspects related to melody, harmony, rhythm, or instrumentation, data-driven approaches often capture confounding factors that may not directly relate to the target concept (e.g., recording conditions in music classification or loudness in singing voice detection).

One main advantage of classical knowledge-based engineering approaches is that they result in explainable and explicit models that can be adjusted intuitively. On the downside, such hand-engineered approaches not only require profound signal processing skills as well as domain knowledge, but also may result in highly specialized solutions that cannot be directly transferred to other problems.

As mentioned earlier, one strong advantage of deep learning is its ability to learn, rather than hand-design, features as part of a model. Nowadays, it seems that attaining state-of-the-art solutions via machine learning depends more on the availability of large quantities of data rather than the sophistication of the approach itself. In this seminar, we critically questioned this statement in the context of concrete music analysis and processing applications. In particular, we explored existing approaches and new directions for combining recent deep learning approaches with classical model-based strategies by integrating knowledge at various stages in the processing pipeline.

There are various ways how one may integrate prior knowledge in DL-based MIR systems. First, one may exploit knowledge already at the input level by using data representations to better isolate information known to be relevant to a task and remove information known to be irrelevant (e.g., by performing vocal source separation before transcribing lyrics). Next, one may incorporate musical knowledge via the model architecture in order to force the model to use its capacity to characterize a particular aspect (e.g., limited receptive fields to prevent a model from “seeing” too much or introducing constraints that mimic DSP systems). Furthermore, the hidden representations can be conditioned to provide humans with “musically sensible control knobs” of the model (e.g., transforming an embedding space to separate out different musical instruments). Knowledge can also be exploited in the design of the output representation (e.g., structured output spaces for chord recognition that account for bass, root, and chroma) or the loss function used for optimization. During the data generation and training process, one may use musically informed data augmentations techniques to enforce certain invariances (e.g., applying pitch shifting to become invariant to musical modulations). Exploiting musical knowledge by combining deep learning techniques with ideas from classical model-based approaches was a core topic of this seminar.

The success of deep learning approaches for learning hidden structures and relations very much depends on the availability of (suitably annotated and structured) data. Therefore, as one fundamental topic, we discussed aspects of generating, collecting, accessing, representing, annotating, preprocessing, and structuring music-related data. These issues are by far not trivial. First of all, music offers a wide range of data types and formats, including text, symbolic data, audio, image, and video. For example, music can be represented as printed sheet music (image domain), encoded as MIDI or MusicXML files (symbolic domain), and played back as audio recordings (acoustic domain). Then, depending on the MIR task, one may need to deal with various types of annotations, including lyrics, chords, guitar tabs, tapping (beat, measure) positions, album covers, as well as a variety of user-generated tags and other types of metadata. To algorithmically exploit the wealth of these various types of information, one requires methods for linking semantically related data sources (e.g., songs and lyrics, sheet music and recorded performances, lead sheet and guitar tabs). Temporal alignment approaches are particularly important to obtain labels for automatic music transcription and analysis tasks. As for data accessibility, copyright issues are the main obstacle for distributing and using music collections in academic research. The generation of freely accessible music (including music composition, performance, and production) requires considerable effort, experience, time, and cost.

Besides the quantity of raw music data and its availability, another crucial issue is the input representation used as the front-end of deep neural networks. For example, log-frequency or Mel spectrograms are often used as input representations when dealing with music signals. We discussed recent research efforts where one tries to directly start with the raw waveform-based audio signal rather than relying on hand-engineered audio representations that exploit domain knowledge. In this context, we discussed how one might resolve phase shift issues by using carefully designed neural network architectures. Further recent research directions include the design of network layers to mimic common front-end transforms or incorporate differentiable filter design methods into a neural network pipeline.

Another central topic we discussed during our seminar was how to exploit musical structures via *self-supervised* and *semi-supervised learning*. Instead of relying on large amounts of labeled data, this technique exploits known variants and invariants of a dataset, using lots of *unlabeled* data. For example, without knowing the transcription of a musical piece, we know how the transcription would change if we shift the whole audio signal by some number of semitones. As another example, we can learn a notion of audio similarity by exploiting the fact that samples from a single musical audio signal are more similar than two samples drawn from different musical audio signals. We also discussed using multi-modal data to give implicit labels, such as text, image, video, and audio correspondences. On the semi-supervised learning side, representations learned in a self-supervised way can be fine-tuned to a particular task with a small amount of labeled data. In this vein, we discussed model generalization, model adaptability, active learning, few-shot learning, and human-in-the-loop systems.

Finally, we addressed topics related to the evaluation of MIR systems. In particular, we discussed the gap between loss functions typically used for optimizing deep learning pipelines and evaluation metrics designed for evaluating specific MIR tasks. In this context, we pointed out the vulnerability of standard metrics to slight variances irrelevant to the perceived output quality, expressing the need for more reliable evaluation metrics. Furthermore, we envisioned the possibility of closing the gap by designing more meaningful loss functions that may be used in the context of end-to-end learning systems.

Participants and Group Composition

In our seminar, we had 22 participants, who came from various locations around the world, including North America (2 participants from the United States), Asia (2 participants from South Korea), and Europe (18 participants from France, Germany, Netherlands, Sweden, United Kingdom). The number of participants and international constellation are remarkable considering the ongoing pandemic. (Note that many of the invited participants, particularly from overseas, were not allowed to go on business trips.) More than half of the participants (12 out of 22) came to Dagstuhl for the first time and expressed enthusiasm about the open and retreat-like atmosphere. Besides its international character, the seminar was also highly interdisciplinary. While most of the participating researchers are working in music information retrieval, we also had participants with a background in musicology, signal processing, machine learning, mathematics, computer vision, and other fields. Our seminar stimulated cross-disciplinary discussions by having experts working in technical and non-technical disciplines while highlighting opportunities for new collaborations among our attendees. Furthermore, the number of participants from the industry (6 out of 22) was relatively high, which also underlines the relevance of the seminar's topic beyond fundamental

research. Most of the participants had a strong musical background, some of them even having a dual career in an engineering discipline and music. This led to numerous social activities, including playing music together. In addition to geographical locations and research disciplines, we tried to foster variety in terms of seniority levels (e.g., we had three Ph.D. students and six participants on the postdoc/junior/assistant professor level) and in terms of gender (6 out of 22 of the participants identify as female). Besides scientific questions, we discussed in our seminar also various challenges that younger colleagues typically face when setting up their research groups and scientific curriculum at the beginning of their academic careers.

Overall Organization and Schedule

Dagstuhl Seminars have a high degree of flexibility and interactivity, allowing participants to discuss ideas and raise questions rather than presenting research results. Following this tradition, we fixed the schedule during the seminar asking for spontaneous contributions with future-oriented content, thus avoiding a conference-like atmosphere, where the focus tends to be on past research achievements. After the organizers gave an overview of the Dagstuhl concept, we started the first day with self-introductions, where all participants introduced themselves and expressed their expectations and wishes for the seminar. We then continued with short (15 to 20 minutes) stimulus talks, where specific participants addressed some critical questions related to the seminar's overall topic in a non-technical fashion. Each of these talks seamlessly moved towards an open discussion among all participants, where the respective presenters took over the role of a moderator. These discussions were well received and often lasted for more than half an hour. The first day closed with a brainstorming session on central topics covering the participants' interests while shaping the overall schedule and format for the next day. We continued having stimulus tasks interleaved with extensive discussions on the subsequent days. On the second day, we split into smaller groups, each group discussing a more specific topic in greater depth. The results and conclusions of these parallel group sessions, which lasted between 60 to 90 minutes, were then presented and discussed with the plenum. However, since the overall seminar size of 22 participants was relatively small, it turned out that the division into subgroups was not necessary. Thanks to excellent group dynamics and a fair distribution of speaking time, all participants had their say and were able to express their thoughts in the plenum while avoiding a monotonous conference-like presentation format. On the last day, we enjoyed a tutorial by Umut Simsekli on some theoretical concepts behind deep learning (a topic unanimously desired by the group). We concluded the seminar with a session we called "self-outroductions" where each participant presented their personal view on the seminar's results.

While working in technical engineering disciplines, most participants also have a strong background and interest in music. This versatility significantly impacted the seminar's atmosphere, leading to cross-disciplinary intersections and provoking discussions and resulting in intensive joint music-making during the breaks and in the evenings. One particular highlight was a concert on Thursday evening organized by Cynthia Liem and Christof Weiß, where various participant-based ensembles performed a wide variety of music, including classical music, Irish folk music, and jazz.

Conclusions and Acknowledgment

There is a growing trend toward building more interpretable deep learning systems, from the data collection and generation stage, to the input and output representations, to the model structure itself. On the other hand, classical model-based approaches bring a wealth of expertise on techniques for knowledge integration in system design. The Dagstuhl Seminar gave us the opportunity for connecting experts from classical model-based approaches, deep learning-based approaches, and related interdisciplinary fields such as music perception and human-computer interaction in order to generate discussion and spark new collaborations. The generation of novel, technically oriented scientific contributions was not the main focus of the seminar. Naturally, many of the contributions and discussions were on a conceptual level, laying the foundations for future projects and collaborations. Thus, the main impact of the seminar is likely to take place in the medium and long term. Some more immediate results, such as plans to share research data and software, also arose from the discussions. As further measurable outputs from the seminar, we expect to see several joint papers and applications for funding.

Besides the scientific aspect, the social aspect of our seminar was just as important. We had an interdisciplinary, international, and interactive group of researchers, consisting of leaders and future leaders in our field. Many of our participants were visiting Dagstuhl for the first time and enthusiastically praised the open and inspiring setting. The group dynamics were excellent, with many personal exchanges and shared activities. Some scientists expressed their appreciation for having the opportunity for prolonged discussions with researchers from neighboring research fields, which is often impossible during conference-like events. At this point, we would like to let some of the participants have their say:

- Stefan Balke (pmOne – Paderborn, DE): *“Dagstuhl is always a wonderful experience, having time to think, talk, and play music. All in a relaxed atmosphere, the seminar feels like a family meeting – especially in these times.”*
- Alice Cohen-Hadria (IRCAM – Paris, FR): *“Now I feel like a part of a community.”*
- Dasaem Jeong (Sogang University – Seoul, KR): *“Full of insightful discussions, music, and friends in a beautiful place.”*
- Cynthia Liem (TU Delft, NL): *“Dagstuhl is the one place in the world where one effectively can have a week long unconference. More deeply talking about research and new ideas, enjoying time with academic friends, with much less distraction than one would have at home, or even in a ‘regular’ conference. Especially coming out of a pandemic, I am realizing this is among the most valuable things in our professional life.”*
- Daniel Stoller (Spotify – Bonn, DE): *“Dagstuhl brings perspectives on the big issues.”*
- Yu Wang (New York University – Brooklyn, US): *“Discussion is like music: the live version is always better.”*

In conclusion, our expectations for the seminar were not only met but exceeded, in particular concerning networking and community building. We want to express our gratitude to the Dagstuhl board for giving us the opportunity to organize this seminar, the Dagstuhl office for their exceptional support in the organization process, and the entire Dagstuhl staff for their excellent service during the seminar. In particular, we want to thank Susanne Bach-Bernhard and Michael Gerke for their assistance during the preparation and organization of the seminar.

2 Table of Contents

Executive Summary

Meinard Müller, Rachel Bittner, and Juhan Nam 104

Stimulus Talks and Further Topics

A Perspective for Machine Learning Education from (Non-Musical) Data Science
Projects
Stefan Balke 111

Toward a New Generation of Source Separation Metrics
Rachel Bittner 111

What is the Best Task to Learn a Generic Music Audio Representation?
Simon Durand and Daniel Stoller 112

What is Actually our Problem? Generalization, Causality, and Error Surfaces
Sebastian Ewert 114

Closing the Gap Between Models and Users
Magdalena Fuentes 115

Performance in Audio/MIDI/Features: Do we Need to Model Underlying Features?
Dasaem Jeong 115

Soft Alignment Techniques for Music Information Retrieval
Michael Krause and Meinard Müller 117

Towards Detecting Musical Patterns in Audio Recordings: A Study on Leitmotifs
Michael Krause, Meinard Müller, and Christof Weiß 118

What Does it Mean to “Work as Intended”?
Cynthia Liem 118

How can we jointly represent the global and local aspects of music at multiple
hierarchical levels?
Gabriel Meseguer Brocal 119

Combining NMF-Based Decomposition and Neural Network Techniques
Meinard Müller and Yigitcan Özer 120

What is the Future of Audio Representations for Music?
Juhan Nam 121

Source Separation of Piano Concertos with Test-Time Adaptation
Yigitcan Özer and Meinard Müller 122

Designing Audio-Specific Deep Learning Front-End Transforms
Geoffroy Peeters 123

Unsupervised Source Separation with Model-Based Deep Learning
Gaël Richard 124

Fractal Structure and Generalization Properties of Stochastic Optimization Al-
gorithms
Umut Şimşekli 125

Thinking Deeply about Ethics
Bob Sturm 126

Knowledge Exchange between Computer Vision and MIR fields <i>Gül Varol</i>	126
Towards Adaptive and Interactive Machine Listening with Minimal Supervision <i>Yu Wang</i>	127
Is Deeper Better? Towards Reliable Evaluation in Music Transcription <i>Christof Weiß and Geoffroy Peeters</i>	128
Working Groups	
Teaching MIR <i>Participants of Dagstuhl Seminar 22082</i>	129
Integrating Musical Knowledge I <i>Participants of Dagstuhl Seminar 22082</i>	130
Integrating Musical Knowledge II <i>Participants of Dagstuhl Seminar 22082</i>	131
Participants	133

3 Stimulus Talks and Further Topics

3.1 A Perspective for Machine Learning Education from (Non-Musical) Data Science Projects

Stefan Balke (pmOne – Paderborn, DE)

License  Creative Commons BY 4.0 International license
© Stefan Balke

A current trend in the industry is to put a strong focus on data and its potential impact on the business side. However, companies tend to underrate the importance of data quality and integration. In fact, 80% of the project time is often spent on these two issues, while only 20% remain for modeling and other issues. However, the expectations on these models are high. Often, there is little knowledge on the business side regarding the mathematical concepts of optimization or machine learning (ML) in general. In addition, the amount of (annotated) data is often very limited. Deep learning approaches are increasingly used, especially for computer vision tasks where pre-trained models are available. In practice, ensemble or boosting methods (e.g., random forests or gradient boosting machines) are still used as the de facto standard in many other applications. Domain knowledge is often integrated into the model via custom features that resemble the underlying business or production processes.

Besides the aforementioned challenges, it is sometimes not clear whether the data resembles or measures the relevant factors (e.g., a missing temperature sensor in a process-critical position will undoubtedly impact the model's performance). Communicating and explaining these challenges to the customer is often more important than trying out the most recent ML approaches. Educators in this field should be aware that many data science projects currently do not fulfill the business side's high expectations. Many stakeholders in such projects still consider machine learning as some "magic" procedure that can solve any problem. Against this background, students should gain experience (e.g., in a class project) in building a dataset from scratch, including the (cumbersome) annotation process, cleaning the data, and developing a model. This experience will help them develop an intuition on the underlying challenges when using algorithms initially designed and tested under lab conditions (e.g., using benchmark datasets). Finally, students should learn to present the results (especially when they are not as expected) in a structured and transparent manner.

3.2 Toward a New Generation of Source Separation Metrics

Rachel Bittner (Spotify – Paris, FR)

License  Creative Commons BY 4.0 International license
© Rachel Bittner

The task of source separation has relied on what has become a standard set of metrics for the past 15 years. These metrics have been the basis for evaluating the latest "state-of-the-art," and determining when design choices are good or bad. Nevertheless, many articles have shown that these metrics are limited in how well they correlate with human perception. The metrics themselves have several undesirable properties, such as being extremely sensitive to phase and unstable at the origin. Moving forward, what could the next generation of metrics look like? The old metrics are not all bad – what can we keep or learn from them? How can

we move beyond “one metric to rule them all” and account for a wider variety of applications and properties to measure? How do we move a community that is very used to ranking all algorithms using one metric away from the concept of “best” or “state-of-the-art”?

References

- 1 Emmanuel Vincent, Rémi Gribonval, Cédric Févotte. Performance measurement in blind audio source separation. *IEEE Transactions on Audio, Speech & Language Processing*, 14(4): 1462–1469 (2006)
- 2 Emmanuel Vincent, Hiroshi Sawada, Pau Bofill, Shoji Makino, Justinian P. Rosca. First Stereo Audio Source Separation Evaluation Campaign: Data, Algorithms and Results. *ICA 2007*: 552–559
- 3 Valentin Emiya, Emmanuel Vincent, Niklas Harlander, Volker Hohmann. Subjective and Objective Quality Assessment of Audio Source Separation. *IEEE Transactions on Audio, Speech & Language Processing*, 19(7): 2046–2057 (2011)
- 4 Emmanuel Vincent. Improved Perceptual Metrics for the Evaluation of Audio Source Separation. *LVA/ICA 2012*: 430–437
- 5 Estefanía Cano, Derry FitzGerald, Karlheinz Brandenburg. Evaluation of quality of sound source separation algorithms: Human perception vs quantitative metrics. *EUSIPCO 2016*: 1758–1762

3.3 What is the Best Task to Learn a Generic Music Audio Representation?

Simon Durand (Spotify – Paris, FR) and Daniel Stoller (Spotify GmbH – Berlin, DE)

License  Creative Commons BY 4.0 International license
© Simon Durand and Daniel Stoller

In the last few years, the fields of computer vision and natural language processing have witnessed a new generation of general-purpose models that can be used for a wide variety of downstream applications [1]. These are pre-trained models using a domain-specific pretext task before applying them to the downstream task of interest. The pretext task is usually conceptually simple and does not require side information such as expensive and often strongly biased labels, but powerful in the sense that a comprehensive understanding of the input is needed to perform well. These approaches also enormously benefit from scaling up model expressivity (such as Transformer models) and dataset size, which are enabled by hardware improvements such as Tensor Processing Units (TPUs).

If such an approach was applied to the music domain, we could obtain models that perform many different music audio analysis tasks even when only a few annotations for a task are available. However, music audio is different from language and vision, and we have not yet found a suitable music processing task with the same simplicity and strength as existent in other domains. Therefore, we review some pretext tasks used in existing approaches and discuss which tasks might be suitable for the music domain.

In natural language processing, randomly masking input tokens and asking the model to reconstruct them was used successfully as a pretext task [2]. Here, one needs to understand the context and semantic meaning of a sentence and its words to predict masked tokens. Similarly, in computer vision, we can predict randomly masked image patches [3]. However, it is questionable whether applying this directly to music audio spectrograms would result in an equally powerful pretext task due to the differing nature of spectrograms and images.

Other generative learning methods are promising pretext tasks, as they encourage a model to learn many semantically relevant features in the process. Generative models can be constructed in different ways, one of which is by using an autoregressive approach. Assuming that the input data can be represented as a sequence of elements, the model must predict the next element in the sequence, given the previous ones. In the NLP domain, [4] used a Transformer-based sequence model to predict the next word given the previous ones, which turned out to acquire many features useful for various downstream tasks. Similarly, making a model continue a given music piece could also yield semantically meaningful features – the usefulness of the Jukebox [5] features can be seen as an early example of such an approach. However, it is more difficult to identify where the relevant features from the model need to be extracted since auto-regressive models do not yield an explicit latent representation. This fact is in contrast to reconstruction-based generative models, such as variational auto-encoders (VAEs) and, in particular, VAEs using vector quantization (VQ-VAEs) that were shown to be effective at compressing music signals into short sequences of tokens [6]. Finally, the flow-based approach has shown some promise in computer vision, where the data is mapped to an equally high-dimensional but ordered latent space using a bijective function. Since the input is not compressed in this approach when moving to the latent space, it might be better suited for tasks where input details are relevant but worse for more abstract tasks.

An alternative task, called contrastive self-supervised learning, is to learn to compare different views of the same input such that the representation puts similar inputs closer together [7]. While it can learn a representation invariant to musical properties through data augmentation, it can be counterproductive for tasks dependent on these properties.

Our discussion is grounded on the specificities of music, on the desirable properties we would like a self-supervision task to highlight, and on the most promising musical applications that could come out of this.

References

- 1 Xiao Liu, Fanjin Zhang, Zhenyu Hou, Zhaoyu Wang, Li Mian, Jing Zhang, Jie Tang. Self-supervised Learning: Generative or Contrastive. CoRR abs/2006.08218 (2020)
- 2 Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL-HLT (1) 2019: 4171–4186
- 3 Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, Neil Houlsby. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. ICLR 2021
- 4 Radford, Alec, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever. Language Models are Unsupervised Multitask Learners. OpenAI Blog, 2019
- 5 Prafulla Dhariwal, Heewoo Jun, Christine Payne, Jong Wook Kim, Alec Radford, Ilya Sutskever. Jukebox: A Generative Model for Music. CoRR abs/2005.00341 (2020)
- 6 Aäron van den Oord, Oriol Vinyals, Koray Kavukcuoglu. Neural Discrete Representation Learning. NeurIPS 2017: 6306–6315
- 7 Michael Gutmann, Aapo Hyvärinen. Noise-contrastive estimation: A new estimation principle for unnormalized statistical models. AISTATS 2010: 297–304

3.4 What is Actually our Problem? Generalization, Causality, and Error Surfaces

Sebastian Ewert (Spotify GmbH – Berlin, DE)

License  Creative Commons BY 4.0 International license
© Sebastian Ewert

As outlined in the Dagstuhl Seminar’s description, data-driven approaches tend to employ confounding factors to achieve results. Based on this observation, I raised a variety of questions in my talk.

- When is the use of confounders actually a problem, and when is it just fine?
 - In other words: Let us say a causal relationship between input and output exists, but the underlying information is hard to extract (which is usually the case). At the same time, some other information is easy to extract but only correlates with the output. Depending on the problem, we mix the uncertainty we get from the difficulty of extracting the proper information with the certainty that we can derive proper conclusions once we have that information. On the other hand, we could use correlations (where we often cannot measure how correlated or causal a factor is) and do everything to optimize average performance. If the latter is a problem, why is 95% of the ML community doing just that and humanity finds it useful?
 - One could argue that we might want to exclude confounding factors if we want to build a causal model of (some aspect of) music, where application performance does not matter, but performance is only helpful to compare different causal models to say which explains the data better. One provocative question is whether we, as MIR researchers, are interested in causal models. Do we understand the underlying principles of music theory and sound production practice? In other words, which aspect of music can not be broken down yet into causal components? Is there something about music we have not yet understood?
- In some cases, the use of confounders is why models do not generalize to new data domains (recording conditions, types of music, and so on). From that perspective, let us discuss when the integration of musical knowledge is actually the best solution to increase generalization capabilities and what “best” means in this case.
- To answer the former question, we should look into alternatives to improve generalizability.
 - In this context, we should discuss what adversarial attacks could solve. What if we encourage submissions that take existing methods and code and try to break it in interesting ways. What is the average Euclidean distance when modifying the input until we observe a confusion or perfect result (for classification problems)? What happens with out-of-domain data and unobserved data? Alternatively, could we encourage authors to attack their own systems? For example, the availability of adversarial attacks on speaker ID systems has led to tons of work to make such systems more robust. What can we learn from that?
 - Is there enough work to understand the error surfaces of neural networks? We have regularizers of all sorts, aiming to make the error more Lipschitz in one way or another. What does that mean? One aspect often discussed in this context is flat areas in error, where small parameter changes only lead to small changes in the output. As music is so structured, is there something here to help us characterize and identify such areas and potentially inspire us to create corresponding regularization approaches?

- Differentiable signal processing components have recently spawned some interest in the community. Do we understand how those should be seen and used in the context of confounders? There are also new stochastic tangent methods to estimate a gradient for non-differentiable components. Is this an option?
- Before the neural network hype, Bayes nets and probabilistic modeling were all the rage. Some of these could be seen as causal models. If the goal is to build and verify proper/causal models for aspects of music, should we not simply ditch neural networks and application performance and go back to these approaches? Alternatively, stochastic nets, such as Bayes neural networks, have recently attracted more and more attention in the theoretical ML world. Is this getting good enough that we should investigate if and how such approaches could build a bridge between the two worlds?

3.5 Closing the Gap Between Models and Users

Magdalena Fuentes (NYU – Brooklyn, US)

License © Creative Commons BY 4.0 International license
© Magdalena Fuentes

Machine learning (ML) now touches all of our lives. Such systems recommend which videos to watch, where to go on holiday, and how to get there. Though progress has been astounding, these ML models fall short when it comes to non-standard examples and subjective applications, such as analyzing underrepresented music genres, understanding people with accents or speech impediments, or recognizing faces from underrepresented ethnicities. ML-based models fall short because they are designed and trained to work well on “averages”. In this context, I discussed in my talk how we could take steps to develop the next generation of ML models to close the gap between the models’ output and the users’ need. I argued how we could leverage several recent advancements in self-supervision, representation learning, and multimodal data analysis to move towards more user-centered ML.

3.6 Performance in Audio/MIDI/Features: Do we Need to Model Underlying Features?

Dasaem Jeong (Sogang University – Seoul, KR)

License © Creative Commons BY 4.0 International license
© Dasaem Jeong
Joint work of Dasaem Jeong, Juhan Nam, Taegyun Kwon

In expressive performance modeling or performance style analysis (e.g., for piano music), it is common to use hand-crafted performance features instead of directly using audio or MIDI representations. One of the reasons is that it makes the problem more focused on *how* someone played, and not *what* or in *which* acoustic condition someone played. Researchers who are using performance features believe that these hand-crafted features somehow better capture the musical intention of the performer. Even when the output of a performance is encoded as audio or MIDI, we assume that hidden underlying features (e.g., related to musical tempo) are present. Explaining a musical performance as a function of an underlying musical

tempo assumes that each note's position in time depends on the performer's internal musical tempo. Using this strong assumption, computational systems for expressive performance modeling have succeeded in mimicking human performances [1, 2].

Most recent deep learning (DL) approaches aim to model outputs in an end-to-end fashion. However, there are examples where those approaches fail, such as generating structured musical pieces with WaveNet trained on piano recordings [3]. On the other hand, the success of DDSP [4] with a limited amount of training data is an example of the effectiveness of modeling underlying features of instrumental sound (which is a combination of harmonic oscillations) instead of modeling the final observable output (i.e., waveform samples).

Still, counterexamples show that only modeling observations without explicit assumptions may be enough to model the complex structure of musical sounds and music. OpenAI's Jukebox [5] models music audio without explicitly modeling underlying features, such as notes, chords, or instrumentation. Also, recent research shows that neural networks that were trained to model observations also learned characteristics of underlying features. Examples are the usage of audio representations obtained by Jukebox for retrieval tasks [6], or the usage of 2D GANs to infer 3D depth [7].

There are further examples in DL-based MIR research, such as using explicit modeling of drum sounds for transcription [8], or using MIDI transcriptions instead of audio for composer identification [9]. However, it is still an open question whether it is beneficial to explicitly model underlying features or model end-level observations with larger-scale training data.

References

- 1 Dasaem Jeong, Taegyun Kwon, Yoojin Kim, Kyogu Lee, Juhan Nam, VirtuosoNet: A Hierarchical RNN-based System for Modeling Expressive Piano Performance. *ISMIR 2019*: 908–915
- 2 Gerhard Widmer, Sebastian Flossmann, Maarten Grachten, YQX Plays Chopin. *AI Mag.* 30(3): 35–48 (2009)
- 3 Aäron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew W. Senior, Koray Kavukcuoglu, WaveNet: A Generative Model for Raw Audio. *SSW 2016*: 125
- 4 Jesse H. Engel, Lamtharn Hantrakul, Chenjie Gu, Adam Roberts. DDSP: Differentiable Digital Signal Processing. *ICLR 2020*
- 5 Prafulla Dhariwal, Heewoo Jun, Christine Payne, Jong Wook Kim, Alec Radford, Ilya Sutskever. Jukebox: A Generative Model for Music. *CoRR abs/2005.00341* (2020)
- 6 Rodrigo Castellon, Chris Donahue, Percy Liang. Codified audio language modeling learns useful representations for music information retrieval. *ISMIR 2021*: 88–96
- 7 Xingang Pan, Bo Dai, Ziwei Liu, Chen Change Loy, Ping Luo. Do 2D GANs Know 3D Shape? Unsupervised 3D Shape Reconstruction from 2D Image GANs. *ICLR 2021*
- 8 Keunwoo Choi, Kyunghyun Cho. Deep Unsupervised Drum Transcription. *ISMIR 2019*: 183–191
- 9 Qiuqiang Kong, Keunwoo Choi, Yuxuan Wang. Large-Scale MIDI-based Composer Classification. *CoRR abs/2010.14805* (2020)

3.7 Soft Alignment Techniques for Music Information Retrieval

Michael Krause (Friedrich-Alexander-Universität Erlangen-Nürnberg, DE) and
Meinard Müller (Friedrich-Alexander-Universität Erlangen-Nürnberg, DE)

License © Creative Commons BY 4.0 International license

© Michael Krause and Meinard Müller

Joint work of Michael Krause, Yigitcan Özer, Christian Dittmar, Frank Zalkow, Meinard Müller

The problem of aligning two time series is central to many areas of music information processing (MIR), including music synchronization, lyrics alignment, and music transcription. At the same time, more and more MIR tasks are approached using deep-learning (DL) models that consist of differentiable building blocks. In traditional music processing, the primary technique used by MIR researchers for aligning sequences of music data is *dynamic time warping* (DTW) [1], which finds a minimum cost alignment between two sequences, given appropriate feature representations and a suitable cost measure. However, the standard DTW recursion involves the computation of hard minima and, therefore, is not differentiable everywhere. Recently, this technique has been modified to be fully differentiable by using a “soft” way to compute minima in an approximative fashion [2, 3]. This strategy opens up the possibility of utilizing a soft DTW variant inside DL systems and backpropagate gradients through an alignment.

In this talk, we focused on music synchronization as a motivating scenario to demonstrate how soft alignment techniques might be helpful in MIR. The idea is to replace a classical pipeline based on chroma features and DTW with a DL-based approach using learned features and a soft DTW variant. We discussed several challenges arising from this related to trivial solutions and the efficiency of the gradient computation. The subsequent discussion focused on other application scenarios, in particular lyrics alignment and lyrics-informed source separation [4]. Furthermore, we discussed the relationship between the soft DTW variant and other techniques such as the connectionist temporal classification (CTC) function, the Viterbi algorithm, and the Baum–Welch algorithm. We found that the soft DTW variant allows for a cleaner formulation of the lyrics alignment problem than commonly used approaches such as CTC, but several technical challenges remain to be solved.

References

- 1 Meinard Müller. Fundamentals of Music Processing – Using Python and Jupyter Notebooks. Second Edition, Springer 2021
- 2 Marco Cuturi, Mathieu Blondel. Soft-DTW: a Differentiable Loss Function for Time-Series. ICML 2017: 894–903
- 3 Isma Hadji, Konstantinos G. Derpanis, Allan D. Jepson. Representation Learning via Global Temporal Alignment and Cycle-Consistency. CVPR 2021: 11068–11077
- 4 Kilian Schulze-Forster, Clément S. J. Doire, Gaël Richard, Roland Badeau. Phoneme Level Lyrics Alignment and Text-Informed Singing Voice Separation. IEEE/ACM Transactions on Audio, Speech and Language Processing, 29: 2382–2395 (2021)

3.8 Towards Detecting Musical Patterns in Audio Recordings: A Study on Leitmotifs

*Michael Krause (Friedrich-Alexander-Universität Erlangen-Nürnberg, DE),
Meinard Müller (Friedrich-Alexander-Universität Erlangen-Nürnberg, DE),
Christof Weiß (Friedrich-Alexander-Universität Erlangen-Nürnberg, DE)*

License  Creative Commons BY 4.0 International license
© Michael Krause, Meinard Müller, and Christof Weiß

During the Dagstuhl Seminar, we discussed some of our recent work on automatically detecting musical patterns in audio recordings with a particular focus on leitmotifs. These motifs are specific patterns associated with certain characters, places, items, or feelings occurring in an opera or a movie soundtrack [2]. Detecting such leitmotifs is particularly challenging since their appearance can change substantially throughout a musical work. Based on a dataset of 200 hours of audio with over 50 000 annotated leitmotif instances, we explored in [1] the benefits and limitations of deep-learning techniques for detecting leitmotifs. To investigate the robustness of the trained systems, we tested their sensitivity to different modifications of the input. We found that our deep-learning systems seem to work well in general. However, a closer investigation reveals that they capture confounding factors such as pitch distributions in leitmotif regions rather than identifying characteristic musical properties such as rhythm and melody. Thus, our in-depth analysis demonstrates some challenges that may arise from applying deep-learning approaches for detecting complex musical patterns in audio recordings. In personal conversations with Dagstuhl participants, we discussed how deep learning systems need to be adapted to detect musical patterns robustly. In particular, we found that these systems need to explicitly model the underlying musical structures (such as melody and rhythm) to prevent them from exploiting confounding factors.

References

- 1 Michael Krause, Meinard Müller, Christof Weiß. Towards Leitmotif Activity Detection in Opera Recordings. Transactions of the International Society for Music Information Retrieval, 4(1): 127-140 (2021)
- 2 Matthew Bribitzer-Stull. Understanding the Leitmotif. Cambridge University Press 2015

3.9 What Does it Mean to “Work as Intended”?

Cynthia Liem (TU Delft, NL)

License  Creative Commons BY 4.0 International license
© Cynthia Liem

With Dagstuhl being a good place for joint reflection, I took the audience along in a central question I have increasingly been pondering: What does it mean to “work as intended”? Considering the nature of our academic research and the applied nature of our work, do we succeed in the way we substantiate and communicate in our papers that our contributions are relevant and impactful? Are our contributions indeed relevant and impactful? What would a contribution need for this? In our discussions throughout the Dagstuhl Seminar, these considerations turned up several times when colleagues raised questions on current methodologies and metrics (associated with the question on “what would (not) get someone a paper”).

Looking at topics surrounding deep learning and its applications, I also engaged in discussions on how to teach these topics to our future generations. I have been wondering whether we currently show and teach the proper examples to them and whether we encourage them to “work as intended” in responsible ways in the future. To illustrate this, in a joint talk with Bob Sturm, I illustrated how misconceptions and bad modeling choices (that a naive student may easily make) led to the childcare benefits scandal in The Netherlands. Next to this, I engaged in discussion groups on MIR teaching, where the broad consensus was that music uniquely gives opportunities to really engage with and consciously perceive data. As part of the discussion following my joint presentation with Bob, I finally raised a controversial question: Is the work we present as “science” actually scientific? Methodologically, I think we are mixing aspects of design, science, and engineering while still seemingly upholding “science” as the most important of these three. However, is this really justified? As a group, we did not reach a consensus on this matter, but lively discussions were held that will hopefully extend beyond this seminar.

3.10 How can we jointly represent the global and local aspects of music at multiple hierarchical levels?

Gabriel Meseguer Brocal (Deezer – Paris, FR)

License © Creative Commons BY 4.0 International license
© Gabriel Meseguer Brocal

Integrating musical knowledge into neural networks can be viewed from two perspectives:

- Using our prior knowledge about a task to guide computation (“telling the model *what* to obtain”).
- Identifying our underlying perceptual mechanisms to construct blocks or learning paradigms to guide the model toward knowledge extraction without enforcing a specific output (i.e., “telling the model *how* to obtain it”).

Both views raise challenging questions. In my presentation, I mainly discussed the latter view, even if it is unclear what our knowledge extraction mechanisms are or how they can be transferred to machine learning models. This hodgepodge includes exploring how we codify patterns and repetitions, develop a sense of hierarchical connections and structures, identify timbral spaces and textures, create context-dependent analysis, or enjoy music through a balance between predictability and surprise.

I find the human capacity to distill complex and compact representations to form high-level ideas at several hierarchical levels and our capacity to work with them fascinating. For instance, we can not only associate a new song to an existing artist (a pure classification task) but also recall it in a few instants without any external musical trigger. This capability shows that we have a deep understanding of what that artist is. The ability to develop hierarchical connections from low-level local elements (i.e., how we process information as a part of a whole) is essential to our knowledge distillation. How to mimic this ability when developing “intelligent” models [1, 2] is a fundamental research question. One main challenge is how to induce this part-whole behavior in our systems to let them derive complex but compact high-level musical ideas from local acoustic observations? These ideas open many other stimulating research questions:

- What are the learning paradigms as well as model and task definitions that induce the right invariances needed to create these capsules of knowledge?

- What is the ideal final high-level representation: a single one, a parse tree, or a graph?
- How do we overcome the problem of dynamically allocating model resources (e.g., to account for the number of essential elements and hierarchical levels that will undoubtedly vary from one song to another)?
- Can we distill knowledge by “listening to” the audio several times and take advantage of overfitting?

By finding answers to these questions (and raising many more), we may be able to capture compact but rich representations and better understand knowledge extraction and model behavior.

References

- 1 Anirudh Goyal, Yoshua Bengio. Inductive Biases for Deep Learning of Higher-Level Cognition. CoRR abs/2011.15091 (2020)
- 2 Geoffrey E. Hinton. How to represent part-whole hierarchies in a neural network. CoRR abs/2102.12627 (2021)

3.11 Combining NMF-Based Decomposition and Neural Network Techniques

Meinard Müller (Friedrich-Alexander-Universität Erlangen-Nürnberg, DE) and Yigitcan Özer (Friedrich-Alexander-Universität Erlangen-Nürnberg, DE)

License  Creative Commons BY 4.0 International license
© Meinard Müller and Yigitcan Özer

Nonnegative Matrix Factorization (NMF) is a powerful technique for factorizing, decomposing, and explaining data [4]. Thanks to its nonnegativity constraints and the multiplicative update rules that preserve these constraints in the training stage, it is easy to incorporate additional domain knowledge that guides the factorization to yield interpretable results. On the other hand, deep neural networks (DNNs), which can learn complex non-linear patterns in a hierarchical manner, have become omnipresent thanks to the availability of suitable hardware and software tools. However, deep learning (DL) models are often hard to interpret and control due to the massive number of trainable parameters. In this talk, we reviewed and discussed current research directions that combine the advantages of NMF-based and DL-based learning approaches. To make our discussion more concrete, we considered an audio decomposition application with the objective to decompose a music recording’s magnitude spectrogram into musically meaningful spectral and activation patterns [1]. In this context, we addressed the following questions:

- How can one redraft an NMF-based decomposition as a DNN-based learning problem using autoencoder-like network architectures? Such an approach was originally proposed in [7].
- How can one retain nonnegativity constraints during the optimization process? Such approaches may be based on projected gradient descent methods [5] and other optimization schemes [10].
- How can one incorporate additional (score-based) prior knowledge into DNN-based models? In this context, one may apply ideas from structured regularization [9] and structured dropout [2].
- How can one simulate stacked NMF-like decomposition technique [8] within the DL framework, thus extending traditional flat NMF architectures?

- What is the effect of invertibility constraints of specific layers on the interpretability of the resulting models? How can such conditions be enforced at the training stage? A first approach was described in [3].

By systematically combining and transferring ideas between NMF-based and DL-based learning approaches, our goal was to understand better the interaction between various regularization techniques and DL-based learning procedures while improving the interpretability of DL-based decomposition results. Furthermore, using audio decomposition as a concrete example, we showed how education in machine learning might be supported by considering motivating and tangible music processing applications [6].

References

- 1 Sebastian Ewert, Meinard Müller. Using score-informed constraints for NMF-based source separation. ICASSP 2012: 129–132
- 2 Sebastian Ewert, Mark B. Sandler. Structured dropout for weak label and multi-instance learning and its application to score-informed source separation. ICASSP 2017: 2277–2281
- 3 Rainer Kelz, Gerhard Widmer. Towards Interpretable Polyphonic Transcription with Invertible Neural Networks. ISMIR 2019: 376–383
- 4 Daniel D. Lee, H. Sebastian Seung. Algorithms for Non-negative Matrix Factorization. NIPS 2000: 556–562
- 5 Chih-Jen Lin. Projected Gradient Methods for Nonnegative Matrix Factorization. Neural Computation, 19(10): 2756–2779 (2007)
- 6 Meinard Müller, Brian McFee, Katherine M. Kinnaird. Interactive Learning of Signal Processing Through Music: Making Fourier Analysis Concrete for Students. IEEE Signal Processing Magazine, 38(3): 73–84 (2021)
- 7 Paris Smaragdis, Shrikant Venkataramani. A neural network alternative to non-negative audio models. ICASSP 2017: 86–90
- 8 George Trigeorgis, Konstantinos Bousmalis, Stefanos Zafeiriou, Björn W. Schuller. A Deep Semi-NMF Model for Learning Hidden Representations. ICML 2014: 1692–1700
- 9 Martin J Wainwright. Structured regularizers for high-dimensional problems: Statistical and computational issues. Annual Review of Statistics and Its Application 1:233–253 (2014)
- 10 Scott Wisdom, Thomas Powers, James W. Pitton, Les Atlas. Deep recurrent NMF for speech separation by unfolding iterative thresholding. WASPAA 2017: 254–258

3.12 What is the Future of Audio Representations for Music?

Juhan Nam (KAIST – Daejeon, KR)

License © Creative Commons BY 4.0 International license
© Juhan Nam

The gist of deep learning (DL) is to learn representations from data to better predict the output or better structure the data within an embedding space. These representations can be learned through neural networks as part of an entire processing pipeline for a given task. For many music information retrieval (MIR) tasks, state-of-the-art neural network models still use time–frequency audio representations designed by explicitly exploiting domain knowledge. Examples are perceptually motivated Mel spectrogram or log-frequency spectrograms obtained by applying a musically motivated constant-Q transform. A recent research trend aims to obtain data-driven audio representations directly from raw audio waveforms (e.g., in sparse coding and convolutional neural networks). While these approaches have shown promising

results in, e.g., music classification tasks, they require large-scale datasets to be successful (e.g., one million songs). Furthermore, these methods do not significantly outperform Mel spectrograms, and the learned (convolution) filters are often hard to interpret. As a compromise between hand-designed and fully-learned audio representations, researchers have attempted to learn filters with constraints (e.g., phase invariance) or learn only parameters of pre-designed filter prototypes. While reviewing recent work on this topic, we discussed possible research directions for learning audio representations and their potential for music processing.

References

- 1 Evan Smith, Michael Lewicki. Efficient Audio Coding. *Nature*, 2006
- 2 Sander Dieleman, Benjamin Schrauwen. End-to-end learning for music audio. *ICASSP 2014*: 6964–6968
- 3 Jongpil Lee, Jiyoung Park, Keunhyoung Luke Kim, Juhan Nam. Sample-level Deep Convolutional Neural Networks for Music Auto-tagging Using Raw Waveforms. *CoRR abs/1703.01789* (2017)
- 4 Taejun Kim, Jongpil Lee, Juhan Nam. Comparison and Analysis of SampleCNN Architectures for Audio Classification. *IEEE Journal of Selected Topics in Signal Processing*, 13(2): 285–297 (2019)
- 5 Jordi Pons, Oriol Nieto, Matthew Prockup, Erik M. Schmidt, Andreas F. Ehmann, Xavier Serra. End-to-end Learning for Music Audio Tagging at Scale. *ISMIR 2018*: 637–644
- 6 Minz Won, Andres Ferraro, Dmitry Bogdanov, Xavier Serra. Evaluation of CNN-based Automatic Music Tagging Models. *CoRR abs/2006.00751* (2020)
- 7 Alexei Baevski, Yuhao Zhou, Abdelrahman Mohamed, Michael Auli. wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations. *NeurIPS 2020*
- 8 Minz Won, Sanghyuk Chun, Oriol Nieto, Xavier Serra. Data-Driven Harmonic Filters for Audio Representation Learning. *ICASSP 2020*: 536–540
- 9 Frank Cwitkowitz, Mojtaba Heydari, Zhiyao Duan. Learning Sparse Analytic Filters for Piano Transcription. *CoRR abs/2108.10382* (2021)

3.13 Source Separation of Piano Concertos with Test-Time Adaptation

Yigitcan Özer (Friedrich-Alexander-Universität Erlangen-Nürnberg, DE) and Meinard Müller (Friedrich-Alexander-Universität Erlangen-Nürnberg, DE)

License © Creative Commons BY 4.0 International license
© Yigitcan Özer and Meinard Müller

Music source separation (MSS) aims at decomposing a music recording into constituent sources, such as a lead instrument and the accompaniment. Despite the difficulties in MSS due to the high correlation of musical sources in time and frequency, deep neural networks (DNNs) have led to substantial improvements to accomplish this task [1, 2]. For training supervised machine learning models such as DNNs, isolated sources are required. In the case of popular music, one can exploit open-source datasets which involve multitrack recordings of vocals, bass, and drums. For western classical music, however, isolated sources are generally not available. In this talk, we considered the case of piano concertos, which are composed for a pianist typically accompanied by an orchestra. The lack of multitrack recordings makes training supervised machine learning models for the separation of piano and orchestra challenging. To overcome this problem, we suggest generating artificial training material by randomly mixing sections of the solo piano repertoire (e.g., piano sonatas) and

orchestral pieces without piano (e.g., symphonies) to train state-of-the-art DNN models for MSS. Furthermore, we propose a test-time adaptation (TTA) procedure, which exploits random mixtures of the piano-only and orchestra-only parts in the test data to finetune the separation quality. To this end, one may first train on an artificially generated dataset. Then, the idea is to exploit that piano concertos comprise long piano-only (e.g., in the *Cadenza*) and orchestra-only (e.g., in the *Exposition*) sections. Using these sections, one can generate synthetic mixtures for the test item at the testing stage to adapt the model and enhance the separation. In this context, we discussed the following points:

- Although random mixes are not musically plausible, they already yield an acceptable separation quality. Would it be feasible to generate further data for this task using an orchestra part (e.g., provided by Music Minus One) of piano concertos and mix them with the piano part played by various pianists on various instruments?
- How can one design an additional loss term to better capture the onsets of the piano?
- Can generative models, e.g., GANs, help to reduce the interference between the separated piano and orchestral sources?
- Would it be a good idea to introduce hierarchical instrument classes to the network such as strings, percussion, woodwinds, and so on?
- Would it be sensible to integrate a piano transcription model into the network?

References

- 1 Romain Hennequin, Anis Khelif, Félix Voituret, Manuel Moussallam. Spleeter: A fast and efficient music source separation tool with pre-trained models. *Journal of Open Source Software*, 5(56): 2154 (2020)
- 2 Fabian-Robert Stöter, Stefan Uhlich, Antoine Liutkus, Yuki Mitsufuji. Open-Unmix – A Reference Implementation for Music Source Separation. *Journal of Open Source Software*, 4(41): 1667 (2019)

3.14 Designing Audio-Specific Deep Learning Front-End Transforms

Geoffroy Peeters (Telecom Paris, FR)

License © Creative Commons BY 4.0 International license
© Geoffroy Peeters

Joint work of Félix Mathieu, Thomas Courtat, Gaël Richard, Geoffroy Peeters

Deep learning front-ends (or encoders) refer to the first projections of a deep neural network (DNN) applied directly to the raw audio waveform. Therefore, the front-end is the DNN part that should be the most adapted to the specificities of the audio signal. Usually, the front-end transform consists of a set of 1D-convolutions applied to the waveform, corresponding to a set of temporal basis functions, kernels, or filters. Such basis functions can simply be the cosine and sine kernels of a short-time Fourier transform (STFT) or constant-Q transform (CQT), the modulus of which brings the nice phase-shift-invariance (PSI) property.

Such kernels can also be trained. For example, the authors of [1] propose to train an end-to-end DNN with a 1D-convolution front-end. In particular, the STFT is mimicked by using real-valued kernels with sizes and strides equal to the STFT window length and hop size. However, this strategy fails to reproduce the STFT's PSI property, requiring the learned kernels to account for every possible phase shift. To reduce the number of required phase shifts to be learned, [2] proposes to reduce the length of the kernels to three samples and adopt an approach similar to the VGG-net architecture (i.e., a pyramid of projections with small kernels).

To better understand what the kernels have learned, the modulus of their DFTs is often analyzed (with the hope that those will highlight band-pass filters). To make this explicit, in the SincNet as used in [3], the kernels are parameterized as band-pass filters (represented as the difference between two low-pass Sinc kernels) with learnable parameters. While interesting, SincNet kernels still do not provide PSI. To ensure PSI, [5] extended the SincNet to the Complex-Gabor, which, apart from bringing the PSI, also provides the best time–frequency trade-off. For the same purpose (getting PSI projections), [6] proposed to define the imaginary kernels as the Hilbert transform of the learned real-valued kernels. Recently, Ditter and Gerkmann [4] showed that it is possible to reach good results using an untrained multi-phase gammatone filter bank. As opposed to Fourier-like kernels, the projection is performed with amplitude-modulated kernels.

To gain the benefits of all the previous approaches, we recently proposed extended Hilbert–Bedrosian kernels, where we learn both a modulating and a modulated carrier basis kernel [7]. The proposed transform achieves the PSI property as well as an envelope-shift-invariance property. We show that we can use large temporal kernels (as the Fourier transform does), hence reducing the overall computational cost with state-of-the-art results for ConvTasNet-like architectures.

References

- 1 Sander Dieleman, Benjamin Schrauwen. End-to-end learning for music audio. ICASSP 2014: 6964–6968
- 2 Jongpil Lee, Jiyoung Park, Keunhyoung Luke Kim, Juhan Nam. Sample-level Deep Convolutional Neural Networks for Music Auto-tagging Using Raw Waveforms. CoRR abs/1703.01789 (2017)
- 3 Mirco Ravanelli, Yoshua Bengio. Speaker Recognition from Raw Waveform with SincNet. SLT 2018: 1021–1028
- 4 David Ditter, Timo Gerkmann. A Multi-Phase Gammatone Filterbank for Speech Separation Via Tasnet. ICASSP 2020: 36–40
- 5 Paul-Gauthier Noé, Titouan Parcollet, Mohamed Morchid. CGCNN: Complex Gabor Convolutional Neural Network on Raw Speech. ICASSP 2020: 7724–7728
- 6 Manuel Pariente, Samuele Cornell, Antoine Deleforge, Emmanuel Vincent. Filterbank Design for End-to-end Speech Separation. ICASSP 2020: 6364–6368
- 7 Félix Mathieu, Thomas Courtat, Gaël Richard and Geoffroy Peeters. Phase Shifted Bedrosian Filter Bank: An Interpretable Audio Front-End for Time-Domain Audio Source Separation. ICASSP 2022

3.15 Unsupervised Source Separation with Model-Based Deep Learning

Gaël Richard (Telecom Paris, FR)

License  Creative Commons BY 4.0 International license
© Gaël Richard

Joint work of Kilian Schulze-Forster, Clément S. J. Doire, Gaël Richard, Roland Badeau

Like many other multimedia-related research areas, music audio analysis has rapidly moved towards pure data-driven deep learning models where domain knowledge is deduced from the processed data. Current state-of-the-art supervised deep learning methods for music source separation require an aligned dataset of mixtures with corresponding isolated source signals. Such datasets are difficult and costly to acquire. In a recent and preliminary study [1], we proposed a novel unsupervised model-based deep learning approach to the specific use

case of choir singing separation. In this work, we represent each source by a differentiable parametric source-filter singing voice model. Then, we train the neural network to reconstruct the observed mixture as a sum of the sources by estimating the source models' parameters given their fundamental frequencies. We believe that integrating domain knowledge in the form of audio models into a data-driven method is key to developing efficient and more frugal machine listening systems. After a brief presentation of this work in the Dagstuhl Seminar, we discussed the different research avenues for exploring and extending this concept of model-based deep learning for music audio analysis.

References

- 1 Kilian Schulze-Forster, Clément S. J. Doire, Gaël Richard, Roland Badeau. Unsupervised Audio Source Separation Using Differentiable Parametric Source Models. CoRR abs/2201.09592 (2022)

3.16 Fractal Structure and Generalization Properties of Stochastic Optimization Algorithms

Umut Şimşekli (INRIA – Paris, FR)

License  Creative Commons BY 4.0 International license
© Umut Şimşekli

Joint work of Alexander Camuto, George Deligiannidis, Murat A. Erdogdu, Mert Gürbüzbalaban, Umut Şimşekli, Lingjiong Zhu

Main reference Alexander Camuto, George Deligiannidis, Murat A. Erdogdu, Mert Gürbüzbalaban, Umut Şimşekli, Lingjiong Zhu: “Fractal Structure and Generalization Properties of Stochastic Optimization Algorithms”, in Proc. of the Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual, pp. 18774–18788, 2021.

URL <https://openreview.net/pdf?id=WlkzLjxpYe>

Understanding generalization in deep learning has been one of the major challenges in statistical learning theory over the last decade. While recent work has illustrated that the dataset and the training algorithm must be taken into account to obtain meaningful generalization bounds, it is still theoretically not clear which properties of the data and the algorithm determine the generalization performance. In this talk, we approached this problem from a dynamical systems theory perspective where stochastic optimization algorithms are represented as random iterated function systems (IFS). Well studied in the dynamical systems literature, such IFSs can be shown (under mild assumptions) to be ergodic with an invariant measure that is often supported on sets with a fractal structure. As our main contribution in [1], we prove that the generalization error of a stochastic optimization algorithm can be bounded based on the complexity of the fractal structure that underlies its invariant measure. Leveraging results from dynamical systems theory, we show that the generalization error can be explicitly linked to the choice of the algorithm (e.g., stochastic gradient descent), algorithm hyperparameters (e.g., step size, batch size), and the geometry of the problem (e.g., Hessian of the loss).

References

- 1 Alexander Camuto, George Deligiannidis, Murat A. Erdogdu, Mert Gurbuzbalaban, Umut Şimşekli, Lingjiong Zhu. Fractal structure and generalization properties of stochastic optimization algorithms. Advances in Neural Information Processing Systems 34 (2021).

3.17 Thinking Deeply about Ethics

Bob Sturm (KTH Royal Institute of Technology – Stockholm, SE)

License  Creative Commons BY 4.0 International license
© Bob Sturm

Main reference Music at the Frontiers of Artificial Creativity and Criticism
URL <https://musaiclab.wordpress.com/>

In my stimulus talk, I reflected on three deep questions:

- How does my work benefit the world?
- How does my work harm the world?
- How do I know?

These are questions of ethics, and reflecting on them provides reusable insights. I illustrated these with my own personal journey of research in applying artificial intelligence to Irish traditional music. Frictions caused by this research motivated me to think about whether such an application harms the world or even whether it provides a benefit – outside of my own professional success. So, I set about learning the tradition by reading, listening, and playing an instrument with a teacher. I have started participating in the tradition to understand the origin of these frictions. These reflections and personal practice have widened my perspectives and developed a new appreciation. I have enlarged my networks outside the ivory tower and am asking more enriching questions. I have deepened my interaction with the world and have become sensitive to responsible engineering – which is constantly under review.

3.18 Knowledge Exchange between Computer Vision and MIR fields

Gül Varol (ENPC – Marne-la-Vallée, FR)

License  Creative Commons BY 4.0 International license
© Gül Varol

In this stimulus talk, we focused on fostering knowledge exchange between computer vision (CV) and music information retrieval (MIR) research communities. First, there was a discussion on *whether* this could be beneficial by illustrating some example works on CV applications on sequential data, highlighting similar problems, and making analogies with MIR. Next, several interdisciplinary research directions were shown to provide examples for *how* to enable more interaction between the two communities.

Neural network architectures increasingly become general-purpose across data modalities, such as image, video, audio, and text. A recent successful example is the family of attention-based Transformer models [9], typically taking sequential data as input. In our works, we adapted this model for problems such as subtitle text alignment in sign language videos [3], text-to-video retrieval [4], and 3D human motion synthesis from textual descriptions [2, 1]. While these tasks are different, common tools can be applied since the data types for video, text, and 3D motion are all similar in their sequential nature. In computer vision, many open questions remain on how to perform temporal modeling best, on how to deal with long-term sequences, and which architecture to employ (e.g., RNNs, CNNs, Transformers, MLPs). Considering that music and audio data are also temporal, any solutions found in one field can inspire the other. Moreover, there exist shared tasks such as alignment, where, for example, subtitle alignment can borrow ideas from MIR problems such as music-lyrics alignment and vice versa.

This Dagstuhl Seminar specifically focused on knowledge integration in deep neural networks. This talk provided several examples of how our works in CV inject domain knowledge, such as physics constraints in 3D hand-object reconstruction from images [6] and differentiable human body model SMPL [10] as a network layer [2]. This discussion led to making analogies between the differentiable 3D models and Differentiable Digital Signal Processing (DDSP) tools used in MIR. Furthermore, instead of integrating, one can *extract* knowledge from end-to-end black-box approaches. Using the attention mechanism as a way to perform localization in long sequences implicitly is an example for such knowledge extraction [5].

Several research problems were identified to have the potential to encourage collaborations across CV and MIR communities, among which are music-conditioned dance generation [7] and visual piano transcription [8]. We expect that bringing together researchers from the two fields for these collaborative projects can facilitate general knowledge exchange as a side product.

References

- 1 Mathis Petrovich, Michael J. Black, Gül Varol. TEMOS: Generating diverse human motions from textual descriptions. arXiv, 2022.
- 2 Mathis Petrovich, Michael J. Black, Gül Varol. Action-Conditioned 3D Human Motion Synthesis with Transformer VAE. ICCV 2021: 10965–10975
- 3 Hannah Bull, Triantafyllos Afouras, Gül Varol, Samuel Albanie, Liliane Momeni, Andrew Zisserman. Aligning Subtitles in Sign Language Videos. ICCV 2021: 11532–11541
- 4 Max Bain, Arsha Nagrani, Gül Varol, Andrew Zisserman. Frozen in Time: A Joint Video and Image Encoder for End-to-End Retrieval. ICCV 2021: 1708–1718
- 5 Gül Varol, Liliane Momeni, Samuel Albanie, Triantafyllos Afouras, Andrew Zisserman. Read and Attend: Temporal Localisation in Sign Language Videos. CVPR 2021: 16857–16866
- 6 Yana Hasson, Gül Varol, Dimitrios Tzionas, Igor Kalevatykh, Michael J. Black, Ivan Laptev, Cordelia Schmid. Learning Joint Reconstruction of Hands and Manipulated Objects. CVPR 2019: 11807–11816
- 7 Ruilong Li, Shan Yang, David A. Ross, Angjoo Kanazawa. AI Choreographer: Music Conditioned 3D Dance Generation with AIST++. ICCV 2021: 13381–13392
- 8 A. Sophia Koepke, Olivia Wiles, Yael Moses, Andrew Zisserman. Sight to Sound: An End-to-End Approach for Visual Piano Transcription. ICASSP 2020: 1838–1842
- 9 Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, Illia Polosukhin. Attention is All you Need. NIPS 2017: 5998–6008
- 10 Matthew Loper and Naureen Mahmood and Javier Romero and Gerard Pons-Moll, Michael J. Black. SMPL: A skinned multi-person linear model. ACM Transactions on Graphics, 34(6): 248:1–248:16 (2015)

3.19 Towards Adaptive and Interactive Machine Listening with Minimal Supervision

Yu Wang (New York University – Brooklyn, US)

License  Creative Commons BY 4.0 International license
© Yu Wang

Machine listening aims at endowing a machine with the ability to perceive and understand audio signals as humans do. It can be applied to different audio domains such as music, speech, and environmental sounds. Nowadays, approaches based on deep learning have

become mainstream tools and achieved state-of-the-art performance in multiple research areas, including machine listening. A deep learning model that generalizes well needs to be trained on a large amount of labeled data. While it is easy to collect a large amount of audio data, labeling them is very costly and often requires expert knowledge. Therefore, current studies in machine listening often suffer from small datasets, which lead to poor generalizability and small vocabulary. Existing strategies tackling this labeled data scarcity issue often still require a significant amount of human effort [1] or have generalizability issues when applied to real audio [2]. To address this, we propose a new perspective that instead of focusing on collecting more labeled data to train a giant universal model, we aim for a flexible and customizable system that can adapt to different tasks quickly with the help of minimal supervision from human input. We envision a paradigm where the model can learn to recognize a new target sound at inference time in a real-time, on-the-fly fashion, based on just a handful of examples (e.g., five) provided by a human user. To do so, we leverage metric-based few-shot learning techniques [3, 4] to learn a discriminative embedding space that can generate a robust representation for an unseen novel class based on a few examples. We applied the trained few-shot learning models to various tasks in different audio domains including keyword spotting [5], drum transcription [6], large vocabulary audio tagging [7], and musical source separation [8].

References

- 1 Mark Levy. Improving perceptual tempo estimation with crowd-sourced annotations. ISMIR 2011: 317–322
- 2 Justin Salamon, Duncan MacConnell, Mark Cartwright, Peter Li, Juan Pablo Bello. Scaper: A library for soundscape synthesis and augmentation. WASPAA 2017: 344–348
- 3 Jake Snell, Kevin Swersky, Richard S. Zemel. Prototypical Networks for Few-shot Learning. NIPS 2017: 4077–4087
- 4 Spyros Gidaris, Nikos Komodakis. Dynamic Few-Shot Visual Learning Without Forgetting. CVPR 2018: 4367–4375
- 5 Yu Wang, Justin Salamon, Nicholas J. Bryan, Juan Pablo Bello. Few-Shot Sound Event Detection. ICASSP 2020: 81–85
- 6 Yu Wang, Justin Salamon, Mark Cartwright, Nicholas J. Bryan, Juan Pablo Bello. Few-shot Drum Transcription in Polyphonic Music. ISMIR 2020: 117–124
- 7 Yu Wang, Nicholas J. Bryan, Mark Cartwright, Juan Pablo Bello, Justin Salamon. Few-Shot Continual Learning for Audio Classification. ICASSP 2021: 321–325
- 8 Yu Wang, Daniel Stoller, Rachel M. Bittner, and Juan P. Bello. Few-shot musical source separation. ICASSP 2022

3.20 Is Deeper Better? Towards Reliable Evaluation in Music Transcription

*Christof Weiß (Friedrich-Alexander-Universität Erlangen-Nürnberg, DE) and
Geoffroy Peeters (Telecom Paris, FR)*

License © Creative Commons BY 4.0 International license
© Christof Weiß and Geoffroy Peeters

Main reference Christof Weiß, Geoffroy Peeters: “Deep-Learning Architectures for Multi-Pitch Estimation: Towards Reliable Evaluation”, CoRR, Vol. abs/2202.09198, 2022.

URL <https://arxiv.org/abs/2202.09198>

Extracting pitch information from music recordings is a challenging and essential problem in music signal processing. Frame-wise transcription or multi-pitch estimation aims for detecting the simultaneous activity of pitches in polyphonic music recordings and has recently

seen significant improvements thanks to deep-learning techniques, with a variety of proposed network architectures. In a recent study [1], we tested different architectures based on convolutional neural networks, the U-net structure, and self-attention components, also proposing several modifications to those. When comparing variants of these architectures with varying capacity using the MusicNet dataset [2], we observed the following:

- Most architectures yield competitive results, and larger model variants seem to be beneficial.
- These results substantially depend on randomness in parameter initialization, data augmentation, order of training examples, and dropout.
- In our cross-version evaluation, where we exploit that our dataset contains several performances (or versions) of each musical work, the particular choice of the training–test splits has a crucial influence on the results.

Concluding from these observations, we question the claim of superiority (“state-of-the-art”) for particular architectures given only small improvements in many published results. To approach these problems, we suggest to shift the focus from technical design choices (such as network architectures) to the more general aspects of experimental design, validity, and generalization with a particular consideration of music-specific properties in the datasets.

References

- 1 Christof Weiß and Geoffroy Peeters. Deep-Learning Architectures for Multi-Pitch Estimation: Towards Reliable Evaluation. CoRR abs/2202.09198 (2022)
- 2 John Thickstun, Zaïd Harchaoui, Sham M. Kakade. Learning Features of Music from Scratch. CoRR abs/1611.09827 (2016)

4 Working Groups

4.1 Teaching MIR

Participants of Dagstuhl Seminar 22082

License © Creative Commons BY 4.0 International license
© Participants of Dagstuhl Seminar 22082

In this working group, we exchanged ideas about how we teach music processing and what conditions and requirements we find at our home institutes. Working in different departments (e.g., engineering, computer science, music sciences) and countries (e.g., France, Germany, South Korea, Netherlands, Sweden), we reported very different experiences. However, we agreed that music is a beautiful and instructive application domain, yielding an intuitive entry point to support education in various fields and on various levels [1]. We summarize some of our thoughts and discussion points in more detail below.

- From our practice as teachers, a shared experience is that our students often have diverse backgrounds and needs. Some students come from the music field with a good background in music theory and performance, while others come from more technical disciplines such as computer science or signal processing with solid programming skills. The heterogeneity of the audience within a lecture or course often makes it challenging to reconcile the needs of all participants. On the positive side, new possibilities for interdisciplinary and interactive collaborations (e.g., within the framework of project groups) are created.

- Especially when teaching a mixed audience, it is essential as a teacher to ensure that all students can follow the course. Most of us agreed that a “task-driven” teaching approach might be helpful, where one uses a concrete music processing task (e.g., beat tracking) as a starting point to motivate abstract concepts (e.g., periodicity analysis for time series). Such an approach also makes it possible to discuss music-theoretical concepts and properties of music signals before the task is mathematically modeled and tackled with algorithmic methods.
- We also discussed intensively whether and to what extent current deep learning (DL) techniques should be part of a lecture on music information retrieval (MIR) and music processing. On the one hand, DL-based methods are state-of-the-art for many MIR tasks, which speaks in favor of introducing students to such techniques at an early stage. On the other hand, these techniques are often like black boxes and may not provide profound insights into the MIR task and the underlying music data. We agreed that it is essential for students to gain the ability to question things and develop skills that go beyond “pushing buttons” and applying “off-the-shelf”-methods. As teachers, it is our responsibility to find a good balance to let our students acquire technical skills and gain a deep understanding on what they do.
- An essential aspect of MIR education is understanding the actual music data and its specific properties. Especially in the age of deep learning, a dataset’s composition and data quality are crucial. Knowing possible data biases and annotation ambiguities is essential for understanding the behavior of a DL-based approach. Therefore, to sensitize our students to such aspects, we believe they should create and annotate a dataset themselves during their scientific training.

In conclusion, we agreed that music is an instructive, challenging, and multi-faceted domain of application, which may serve as a motivating and intuitive entry point for teaching and learning a wide range of topics beyond MIR, including signal processing, machine learning, and information retrieval. In addition, teaching a music processing or MIR class also provides an excellent opportunity to address sociological and ethical issues – aspects that are often neglected in the thicket of technological details.

References

- 1 Meinard Müller, Brian McFee, Katherine M. Kinnaird. Interactive Learning of Signal Processing Through Music: Making Fourier Analysis Concrete for Students. *IEEE Signal Processing Magazine*, 38(3): 73-84 (2021)

4.2 Integrating Musical Knowledge I

Participants of Dagstuhl Seminar 22082

License  Creative Commons BY 4.0 International license
© Participants of Dagstuhl Seminar 22082

In our group discussion, we identified a multitude of ways how musical knowledge could be incorporated into deep learning (DL) models. Almost all system design choices are, or can be informed by understanding the problem domain, including training data, augmentations, invariances, conditioning, representations, features, model architectures, loss functions and evaluation metrics. We then debated whether knowledge integration is a useful or necessary strategy, and how its influence and success could be measured. We noted the difficulty of discussing the topic without reference to one or more specific tasks. Unlike in other fields,

where domain knowledge is grounded on physical objects, music is an abstract art form where musical concepts are often hard to grasp. In our discussion, going beyond Western music theory, we attempted to define musical knowledge broadly, covering general (perceptually motivated) concepts such as expectation and surprise. We agreed that musical knowledge can be used as an inductive bias to favour musically likely outcomes, but it remained unclear how one could demonstrate that this bias is more useful (or fair) than the natural biases of the datasets used. Furthermore, we identified example cases where musical knowledge has proven beneficial, and we mentioned HCQT as an input representation, music language models in transcription, and prior knowledge of instrumentation of the input data when performing source separation. We also discussed the potential reduction in training data requirements as an advantage of using domain knowledge, and thus in computational cost. Energy metrics would be one way to demonstrate this advantage, as a contribution to Green AI.

4.3 Integrating Musical Knowledge II

Participants of Dagstuhl Seminar 22082

License  Creative Commons BY 4.0 International license
© Participants of Dagstuhl Seminar 22082

In this working group, we discussed the question of knowledge integration for deep-learning systems in music information retrieval. First, we asked ourselves *where* and *how* to integrate knowledge: typical points of leverage can be the dataset creation, the input representation, the network architecture, the loss, or data augmentation strategies. Moreover, knowledge integration at the front-end side (signal processing knowledge) can have different effects than that at the back-end side (musicological or user knowledge). Having an interpretable intermediate representation can help to understand the behavior for downstream tasks as shown for music auto-tagging [1]. Whether knowledge integration is beneficial depends on the desired application, which can be categorized according to various dimensions such as level of expertise or type of task (generation or analysis). For instance, unexpected behavior of an end-to-end system might be considered valuable (“creative”) for music generation. We then turned to an in-depth discussion about the benefits of knowledge integration instead of end-to-end learning. On the one hand, end-to-end systems have shown to be superior to systems integrating knowledge in fields such as computer vision. Moreover, the question arises whether integrating knowledge always leads to a bias. For instance, an input representation relying on the harmonic series will bias results towards music with harmonic instruments and may not work well with inharmonic instruments such as bells, gongs, or low piano notes. On the other hand, such a bias or characteristic behavior may be desired when targeting specific application scenarios. In this sense, knowledge integration can help make a generation system more controllable or an analysis system more objective, interpretable, or plausible. In particular, an analysis system with a human in the loop can greatly use musical knowledge for the reasons mentioned above. Furthermore, all participants agreed that knowledge integration has a high potential to account for smaller and more efficient networks, which helps to move forward towards “Green AI” without losing performance. Finally, touching upon questions raised in the earlier stimulus talks and panel discussions, we agreed that evaluation of model performance is tricky for the case of music information retrieval. Dataset size, annotation quality, annotator bias, and problematic evaluation metrics

often limit the validity of experiments, especially given the small-scale improvements usually reported for novel systems. If we consider today’s results in various MIR tasks to be close to a certain upper bound or “glass ceiling,” the main value of knowledge integration may be getting to this upper bound in a more resource-efficient and interpretable way.

References

- 1 Minz Won, Sanghyuk Chun, Xavier Serra. Toward Interpretable Music Tagging with Self-Attention. CoRR abs/1906.04972 (2019)

Participants

- Stefan Balke
pmOne – Paderborn, DE
- Rachel Bittner
Spotify – Paris, FR
- Alice Cohen-Hadria
IRCAM – Paris, FR
- Simon Dixon
Queen Mary University of
London, GB
- Simon Durand
Spotify – Paris, FR
- Sebastian Ewert
Spotify GmbH – Berlin, DE
- Magdalena Fuentes
NYU – Brooklyn, US
- Dasaem Jeong
Sogang University – Seoul, KR
- Michael Krause
Friedrich-Alexander-Universität
Erlangen-Nürnberg, DE
- Cynthia Liem
TU Delft, NL
- Gabriel Meseguer Brocal
Deezer – Paris, FR
- Meinard Müller
Friedrich-Alexander-Universität
Erlangen-Nürnberg, DE
- Juhan Nam
KAIST – Daejeon, KR
- Yigitcan Özer
Friedrich-Alexander-Universität
Erlangen-Nürnberg, DE
- Geoffroy Peeters
Telecom Paris, FR
- Gaël Richard
Telecom Paris, FR
- Umut Simsekli
INRIA – Paris, FR
- Daniel Stoller
Spotify GmbH – Berlin, DE
- Bob Sturm
KTH Royal Institute of
Technology – Stockholm, SE
- Gül Varol
ENPC – Marne-la-Vallée, FR
- Yu Wang
New York University –
Brooklyn, US
- Christof Weiß
Friedrich-Alexander-Universität
Erlangen-Nürnberg, DE



AI for the Social Good

Claudia Clopath^{*1}, Ruben De Winne^{*2}, and Tom Schaul^{*3}

1 Imperial College London, GB. c.clopath@imperial.ac.uk

2 Oxfam Novib – The Hague, NL. ruben.dewinne@oxfamnovib.nl

3 Google DeepMind – London, GB. schaul@google.com

Abstract

Progress in the field of Artificial intelligence (AI) and machine learning (ML) has not slowed down in recent years. Long-standing challenges like Go have fallen and the technology has entered daily use via the vision, speech or translation capabilities in billions of smartphones. The pace of research progress shows no signs of slowing down, and demand for talent is unprecedented. AI for Social Good in general is trying to ensure that the social good does not become an afterthought, but that society benefits as a whole. In this Dagstuhl Seminar, which can be considered a follow-up edition of Dagstuhl Seminar 19082, we brought together AI and machine learning researchers with non-governmental organisations (NGOs), as they already pursue a social good goal, have rich domain knowledge, and vast networks with (non-)governmental actors in developing countries. Such collaborations benefit both sides: on the one hand, the new techniques can help with prediction, data analysis, modelling, or decision making. On the other hand, the NGOs' domains contain many non-standard conditions, like missing data, side-effects, or multiple competing objectives, all of which are fascinating research challenges in themselves. And of course, publication impact is substantially enhanced when a method has real-world impact. In this seminar, researchers and practitioners from diverse areas of machine learning joined stakeholders from a range of NGOs to spend a week together. We first pursued an improved understanding of each side's challenges and established a common language, via presentations and discussion groups. Building on this foundation, we organised a hackathon around some existing technical questions within the NGOs to scope the applicability of AI methods and seed collaborations. Finally, we defined guidelines and next steps for future AI for Social Good initiatives.

Seminar February 27–4, 2022 – <http://www.dagstuhl.de/22091>

2012 ACM Subject Classification Computing methodologies → Machine learning

Keywords and phrases Machine Learning, Artificial Intelligence, Social Good, NGO, sustainable development goals, Non-governmental organisation

Digital Object Identifier 10.4230/DagRep.12.2.134

1 Executive Summary

Ruben De Winne (Oxfam Novib – The Hague, NL, ruben.dewinne@oxfamnovib.nl)

License  Creative Commons BY 4.0 International license
© Ruben De Winne

AI and ML have made impressive progress in the last few years. Long-standing challenges like Go have fallen and the technology has entered daily use via the vision, speech or translation capabilities in billions of smartphones. The pace of research progress shows no signs of slowing down, and demand for talent is unprecedented. But as part of a wider AI for Social Good trend, this seminar wanted to contribute to ensuring that the social good does not become an afterthought in the rapid AI and ML evolution, but that society benefits as a whole. The

* Editor / Organizer

five-day seminar brought together AI and ML researchers from various universities with representatives from NGOs based in Benin, Tanzania, Uganda, The Netherlands and globally. These NGOs all pursue various social good goals, such as improving air quality, increasing agricultural productivity with the help of technology, transforming health care, providing humanitarian support, and defeating poverty. On these topics, NGOs have rich domain knowledge, just like they have vast networks with (non-)governmental actors in developing countries. Mostly, NGOs have their finger on the pulse of the challenges that the world and especially its most vulnerable inhabitants are facing today, and will be facing tomorrow. The objective of the seminar was to look at these challenges through an AI and ML lens, to explore if and how these technologies could help NGOs to address these challenges. The motivation was also that collaborations between AI and ML researchers and NGOs could benefit both sides: on the one hand, the new techniques can help with prediction, data analysis, modelling, or decision making. On the other hand, the NGOs' domains contain many non-standard conditions, like missing data, side-effects, or multiple competing objectives, all of which are fascinating research challenges in themselves. And of course, publication impact is substantially enhanced when a method has real-world impact. The seminar facilitated the exploration of possible collaborations between AI and ML researchers and NGOs through a two-pronged approach. This approach combined high-level talks and discussions on the one hand with a hands-on hackathon on the other hand. High-level talks and discussions focused first on the central concepts and theories in AI and ML and in the NGOs' development work, before diving into specific issues such as generalisability, data pipelines, and explainability. These talks and discussions allowed all participants – in a very short time-frame – to reach a sufficient level of understanding of each other's work. This understanding was the basis to then start investigating jointly through a hackathon how AI and ML could help addressing the real-world challenges presented by the NGOs. At the start of the hackathon, an open marketplace-like setting allowed AI and ML researchers and NGOs to find the best match between technological supply and demand. When teams of researchers and NGOs were established, their initial objective was not to start coding, but to define objectives, assess scope and feasibility. The intense exchanges during the hackathon allowed NGOs with a lower AI/ML maturity increased to increase understanding of the capabilities of AI/ML and define actions to effectively start working with AI/ML. NGOs that already had a more advanced understanding and use of AI/ML technology prior to the seminar, could take their AI maturity to the next level by trying out new ML approaches, designing and testing tailored ML models, or simply exploring new partnerships. Key to this success of the hackathon – and the seminar at large – was the presence of AI/ML experts whose respective fields of expertise could seamlessly be matched with the various needs of the various NGOs. This excellent group composition also facilitated a productive discussion about guidelines on how to do effective AI for social good collaborations in the future (e.g. by focusing on long-term partnerships, and by sequencing problem scoping before data cleaning and – only in last instance – an actual hackathon).

2 Table of Contents

Executive Summary	
<i>Ruben De Winne</i>	134
Overview of Talks	
Introduction to the seminar	
<i>Claudia Clopath</i>	137
Introduction to Artificial Intelligence	
<i>Tom Schaul</i>	137
Deploying computer vision for conservation	
<i>Sara Beery</i>	138
Dark ecology: unraveling mysteries of bird migration using RADAR and Machine Learning	
<i>Subhransu Maji</i>	138
Working groups	138
Hackathon	139
Guidelines on how to do effective AI for social good collaborations in the future	140
Follow-up actions	141
Participants	142

3 Overview of Talks

3.1 Introduction to the seminar

Claudia Clopath (Imperial College London, GB)

License  Creative Commons BY 4.0 International license
© Claudia Clopath

Claudia Clopath introduced the notion of “AI for Social Good” and the motivation for the seminar. In particular, she listed the following goals: 1) Bring NGO participants and AI researchers together, 2) Create awareness, 3) Make sure incentives are aligned, 4) New innovations coming out of the newly formed collaborations. The potential outcomes of the seminars were: 1) Understanding the link between the two worlds: Machine Learning and NGOs, 2) Understanding the ways to make AI for Social Good initiatives work, 3) Build Prototypes during the hackathon time, 4) Build collaborations, create a network, make friends, 5) and prepare a set of follow-up plans. Finally, she defined a set of rules for the seminar: Connect, Respect, Bottom-Up, Brainstorm, Fun.

3.2 Introduction to Artificial Intelligence

Tom Schaul (Google DeepMind – London, GB)

License  Creative Commons BY 4.0 International license
© Tom Schaul

Tom Schaul gave a brief overview of Artificial Intelligence. He talked about the promise of AI, gave an overview of the subfields of AI terminology, explained some key concepts specifically around the types of machine learning (i.e. supervised learning, unsupervised learning, and reinforcement learning), as well as the most common issues (grouped into data, model and process challenges). An important comment made during the session was that a human needs to check results when deploying machine learning in the real world. In addition, NGO participants in the session were advised that simple but trustable, interpretable and robust methods should always be the first they start with. Only when simple methods aren't sufficient, then one can go to more complex methods such as neural networks. Participants were also warned about the risk of overfitting, which happens when the model becomes so good, doesn't make any mistakes anymore, and gets too precise, and gets bad at anything else. E.g. when detecting a blue pixel in images with cats leads to predicting a “cat” on images with a blue pixel but no actual cat. To mitigate the risk of overfitting, the choice of a validation set is very important: it is advised to use a validation set that uses data from the real-world use case the NGO is working with.

3.3 Deploying computer vision for conservation

Sara Beery (California Institute of Technology – Pasadena, US)

License  Creative Commons BY 4.0 International license
© Sara Beery

Sara Beery presented a use case of AI for social good, i.e. the deployment of computer vision for ecological conservation. She addressed challenges in deploying impactful computer vision solutions for problems that are faced by conservation land managers, noting the importance of also being honest about failures. She pointed to usefulness, accessibility, collaborative development and feedback, and strong communication of risks and intended use as the main factors that enable impact of computer vision in the field of ecological conservation. In addition, interdisciplinary collaboration was mentioned as a factor leading to deeper understanding and new interesting challenges for computer vision. To conclude, Sara emphasized the need for a sustainable plan for supporting any tools that were developed. Indeed, long-term support is needed, as a computer vision model in itself is not a solution. For NGOs, it is critical to empower them by building capacity within the community and/or by advocating for the community to hire the right capacity.

3.4 Dark ecology: unraveling mysteries of bird migration using RADAR and Machine Learning

Subhransu Maji (University of Massachusetts – Amherst, US)

License  Creative Commons BY 4.0 International license
© Subhransu Maji

Subhransu Maji started his presentation by explaining that billions of birds migrate every year, mostly in the cover of darkness. But these birds are visible on RADAR networks in the continental US. Subhransu explained that using Machine Learning, we can learn how migration has changed over the last 25+ years. A team of ecologists and computer scientists worked together to analyze this bird migration data at scale. Challenges and unique opportunities that this collaboration had were also discussed. Subhransu concluded with three pieces of advice to the participants:

- Don't throw away noisy data! You might be able to correct for noise.
- Don't throw away info on who labeled data!
- Don't throw away intermediate things, might be useful at some point for training.

4 Working groups

The seminar agenda left ample room for spontaneous discussions. A brainstorm session led participants to select the following topics for further exploration in smaller working groups: Off-the-shelf AI apps, tools, and platforms, and what holds NGOs back from using these?

- Task-generalizability and regional generalizability
- Setting up data pipelines connected to model deployment, i.e. getting data, labeling data, how to get data pipelines up and running
- What methods exist to understand text
- Tools to support explainability and decision-making
- ML project management: feedback, sustainability, incentives, defining the problem

At the end of these discussions, the groups came back together in the plenary session and shared their main conclusions with each other. The group that had discussed generalizability for instance, noted that generalizability of your model should never be assumed, but should always be tested. The group that discussed data pipelines fed back to the wider group that certain data sets are made available, e.g. by governments/government-funded agencies, and that dealing with anomalies is important, which is especially true for sensor data (because sensors sometimes break etc.).

5 Hackathon

During the hackathon, which was spread out over two days, five groups consisting of machine learning experts and NGOs tackled the real-world issues that the latter had brought to Dagstuhl. As several NGOs had brought issues that required similar technological expertise, they joined forces and benefited from each other's perspectives, challenges, and lessons learned.

Group 1 AirQo from Uganda and Laterite from the Netherlands brought seemingly different cases to the table, i.e. measuring air quality more efficiently with a limited number of sensors placed at various locations across Uganda (fixed locations as well as in moving vehicles), and predicting school drop-out on the basis of various sources of data, for example, surveys of populations, respectively. Nonetheless, in machine learning terms what unified both cases was that their data was feature-based (and not, for example, image- or text-based). Machine Learning algorithms such as Gaussian Processes (GPs) and Gradient Boosted Decision Trees (GBDTs) perform well for such use-cases. Before actually applying these algorithms to their respective cases, both AirQo and Laterite needed to rephrase their problem and properly set up their data pipeline. Laterite went with using GBDTs while AirQo used GPs to balance various trade-offs of these algorithms. At the end of the hackathon, different approaches had been tested with real-world data, and the group had agreed to continue their collaboration beyond the seminar.

Group 2 Soon dubbed “the text group” – Oxfam Novib, Save the Children, and the Red Cross joined forces to discuss and address challenges with natural language. While Oxfam Novib and Save the Children aimed at automating knowledge management and reporting, the Red Cross wanted an algorithm to classify open-text survey responses. Oxfam Novib and Save the Children invested their time mostly in scoping their problem, conducting exploratory online interviews with internal stakeholders, and designing a plan for action after the seminar. The Red Cross sub-team did develop a prototype algorithm that could effectively do the desired classification. The team also considered applicability of the model beyond the particular domain it looked at for this case (i.e. rumors and opinions about COVID-19). It documented the results of tests with various models. And last but not least, it secured internal funding at the Red Cross to continue working on this project.

Group 3 TechnoServe, Humanitarian OpenStreetMap, and again the Red Cross all brought cases requiring computer vision technology. This group probably got furthest with developing and testing actual machine learning models. At the end of the hackathon, their model could indeed recognize trees (catering to TechnoServe's case) and recognize buildings (useful for the cases of Humanitarian OpenStreetMap and the Red Cross). The machine learning experts in the group committed to remaining available for questions after the seminar, and the NGOs would give a demonstration of the models in their

respective organizations.

Group 4 In the fourth and last group, D-tree from Tanzania, which has strong in-house machine learning expertise, mostly scoped out ideas of how to further advance their use of machine learning. Together with the machine learning experts at the seminar, they did two iterations on these ideas with the team in Tanzania. This resulted in a focused list of ideas that D-tree could pick up after the seminar (e.g. solving the interpretability problem, tailoring their health survey, designing a more continuous instead of binary scale, personalizing incentives, detecting suspicious visits, and predicting the type of intervention instead of predicting the outcome). For most of these ideas, simple solutions such as software engineering could already be very valuable. To bring these ideas into practice, the group also identified possible collaborations and partnerships in this area. All in all, the hackathon was a success: NGOs with a lower AI/ML maturity increased their understanding of the capabilities of AI/ML, while NGOs that already had a more advanced understanding and use of AI/ML technology could take a next step. Key to this success was the presence of AI/ML experts whose respective fields of expertise could seamlessly be matched with the various needs of the various NGOs. In times of COVID-19 – with the many impediments to international travel and even a fair share last-minute cancellations – the seminar was fortunate to have this nearly perfect match between offer and demand of skills and expertise.

6 Guidelines on how to do effective AI for social good collaborations in the future

On the final day of the seminar, the participants reflected on the success of the seminar and on what had been the enabling factors. They formulated a set of guidelines on how to do effective AI for social good collaborations in the future. For seminars like the AI for Social Good seminar at Dagstuhl, the importance of in-person attendance was underscored. In addition to AI/ML researchers and domain experts, software engineers should be invited, and the presence of financial partners could be considered. Also, the seminar programme could include a session on what AI can and cannot do, if several NGO participants lack this knowledge. For AI/ML participants, a session or a talk at the start of the seminar about structures and constraints for NGOs could be useful. In particular, domain experts should “educate” AI experts on what is needed/feasible in the field, so that expectations of both worlds can be better aligned. Concretely: AI experts should share experiences from previous (failed) AI for social good pilots/experiments/projects. For visibility purposes, it’s advised to broadly advertise AI for Social Good seminars to NGOs (e.g. through NetHope, or at individual organizations’ events) and to invest in communication via social media (e.g. by publishing success stories of seminar 22091 to justify NGOs’ time). In general, AI for social good collaborations should be clear from the outset about the willingness on both sides to pursue the collaboration in the long-term. Long-term partnerships should be established between affiliations/organizations, not (only) between individuals. But attention point: partnerships between organizations can be painful to set up. To facilitate long-term partnerships, a GitHub team page/repository could be set up with a to-do list per project, so that others can pick up work when they have time, or at another workshop etc. A concrete avenue here would be to tap into existing summer schools or other programmes of universities related to AI for social good (e.g. UMass Amherst, University of Washington, DSSG, Data

summer school at Ecole polytechnique in France, DeepLearning Indaba). More long-term partnerships with academics can also be very useful, because they have lots of students who need lots of projects. The challenge here is that the students are junior, so the senior academic needs to buy into the collaboration and supervise their students.

AI for social good collaborations should ideally follow the following sequence: 1) a scoping exercise to clearly define the problem and identify potentially relevant datasets; 2) a “cleanathon” where software engineers look at the data, 3) the actual hackathon(s) where algorithms/models are designed and tested. In-person interactions are deemed effective for all three of these phases.

Finally, it was suggested to bring NGOs to AI conferences and workshops (e.g. the likes of NeurIPS, etc.).

7 Follow-up actions

To close the seminar while opening concrete perspectives for longer term collaboration, the four groups agreed on action plans with the following structure:

- What to do next week
- What to do next month
- What to do next year
- Who does what?

The entire group also agreed to explore collaboration with HuggingFace, which could host AI for social good models and provide additional (paid) services. They also agreed to have an (online) follow-up session in September 2022, during which they would honestly share experiences with each other and pitch the results achieved at the Dagstuhl Seminar to funders.

Participants

- Emran Alchikh Alnajar
Humanitarian OpenStreetMap,
DE
- Asma Atamna
Ruhr-Universität Bochum, DE
- Sara Beery
California Institute of Technology
– Pasadena, US
- Claudia Clopath
Imperial College London, GB
- Ruben De Winne
Oxfam Novib – The Hague, NL
- Jose Pablo Folch
Imperial College London, GB
- Jörg Franke
Universität Freiburg, DE
- Mariella Goebel
TechnoServe – Cotonou, BJ
- Nitusima Kataraiia
D-tree – Dar es Salaam, TZ
- Stijn Koster
Save the Children –
The Hague, NL
- Rik Linssen
Laterite – Amsterdam, NL
- Subhransu Maji
University of Massachusetts –
Amherst, US
- Jacopo Margutti
510 / Netherlands Red Cross –
The Hague, NL
- Corina Markodimitraki
510 / Netherlands Red Cross –
The Hague, NL
- Nele Quast
University of Oxford, GB
- Raghu Rajan
Universität Freiburg, DE
- José Ricardo Rubio Valverde
Laterite – Amsterdam, NL
- Samira Said
D-tree – Dar es Salaam, TZ
- Tom Schaul
Google DeepMind – London, GB
- Ruby Sedgwick
Imperial College London, GB
- Richard Sserunjogi
AirQo – Kampala, UG
- Christoph Weigl
TechnoServe – Arlington, US

