DAGSTUHL
REPORTS

Volume 13, Issue 4, April 2023

Dagstuhl Reports, Vol. 13, Issue 4 ISSN 2192-5283

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
November, 2023

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`.

Digital Object Identifier: 10.4230/DagRep.13.4.i

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.

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

Report from Dagstuhl Seminar 23151

Normative Reasoning for AI

Agata Ciabattoni[*1], **John F. Horty**[*2], **Marija Slavkovik**[*3], **Leendert van der Torre**[*4], **and Aleks Knoks**[†5]

1 **TU Wien, AT.** agata@logic.at
2 **University of Maryland – College Park, US.** horty@umd.edu
3 **University of Bergen, NO.** marija.slavkovik@uib.no
4 **University of Luxembourg, LU.** leon.vandertorre@uni.lu
5 **University of Luxembourg, LU.** aleks.knoks@uni.lu

Abstract

Normative reasoning is reasoning about normative matters – such as obligations, permissions, and the rights of individuals or groups. It is prevalent in both legal and ethical discourse, and it can – and arguably should – play a crucial role in the construction of autonomous agents. We often find it important to know whether specific norms apply in a given situation, and to understand why and when they apply, and why some other norms do not apply. In most cases, our reasons for wanting to know are purely practical – we want to make the correct decision – but they can also be more theoretical – as they are when we engage in theoretical ethics. Either way, the same questions are crucial for designing autonomous agents sensitive to legal, ethical, and social norms. This Dagstuhl Seminar brought together experts in computer science, logic (including deontic logic and argumentation), philosophy, ethics, and law with the aim of finding effective ways of formalizing norms and embedding normative reasoning in AI systems. We discussed new ways of using deontic logic and argumentation to provide explanations answering normative why questions, including such questions as "Why should I do A (rather than B)?", "Why should you do A (rather than I)?", "Why do you have the right to do A despite a certain fact or a certain norm?", and "Why does one normative system forbid me to do A, while another one allows it?". We also explored the use of formal methods in combination with sub-symbolic AI (or Machine Learning) with a view towards designing autonomous agents that can follow (legal, ethical, and social) norms.

Seminar April 10–14, 2023 – https://www.dagstuhl.de/23151
2012 ACM Subject Classification Computing methodologies → Artificial intelligence; Theory of computation → Logic; Computing methodologies → Multi-agent systems
Keywords and phrases deontic logic, autonomous agents, AI ethics, deontic explanations
Digital Object Identifier 10.4230/DagRep.13.4.1

* Editor / Organizer
† Editorial Assistant / Collector

Normative Reasoning for AI, *Dagstuhl Reports*, Vol. 13, Issue 4, pp. 1–23
Editors: Agata Ciabattoni, John F. Horty, Marija Slavkovik, Leendert van der Torre, and Aleks Knoks
Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Executive Summary

Agata Ciabattoni (TU Wien, AT)
John F. Horty (University of Maryland – College Park, US)
Marija Slavkovik (University of Bergen, NO)
Leendert van der Torre (University of Luxembourg, LU)

Normative reasoning – or, roughly, reasoning about such normative matters as obligations, permissions, and rights – is receiving increasing attention in several fields related to AI and computer science. There is an increase in its more traditional use in knowledge representation and reasoning, multiagent systems, and AI & law. However, it holds much promise and is also becoming more important in the context of the blooming fields of AI ethics and explainable AI. Accordingly, the interdisciplinary seminar *Normative Reasoning for Artificial Intelligence* brought together researchers working in knowledge representation and reasoning, multiagent systems, AI & law, AI ethics, and explainable AI to discuss ways in which normative reasoning can be used to make progress in the latter two disciplines.

While this Dagstuhl Seminar touched upon many different aspects of normative reasoning in AI, four topics received particular attention: (i) from AI & law to AI ethics, (ii) deontic explanations, (iii) defeasible deontic logic and formal argumentation, and (iv) from theory to tools.

From AI & law to AI ethics. AI & law is a field that is concerned with, on the one hand, laws that regulate the use and development of artificial intelligence and, on the other, the use of AI by lawyers and the impact of AI on the legal profession. In this field, normative systems are often used to represent and reason about the legal code. The seminar participants explored different ways in which ideas from AI & law can be used in the context of AI ethics.

Deontic explanations. This topic had to do with the use of formal methods, in general, and deontic logic and the theory of normative systems, in particular, to provide answers to why questions involving deontic expressions: "Why must I wear a face mask?", "Why is it forbidden for me to go out at night, although that other person is allowed to go out at night?", "Why has the law of privacy been changed in this way?". Deontic explanations have an essentially practical nature, which distinguishes them from (merely) scientific explanations. The concerns of scientific explanations focus on causality and uncertainty, whereas deontic explanations additionally include preferences, norms, sanctions, and actions. While causality and uncertainty are core concerns in explainable AI, in the context of our seminar, they played a relatively minor role. Instead, the seminar focused on the aspects of deontic explanations that are special to deontic explanations.

Defeasible deontic logic and formal argumentation. The third topic of the seminar had to do with the role of nonmonotonicity in deontic logic in general and the use of formal argumentation in particular. As is well known in the area of deontic logic, normative reasoning comes with its own set of benchmark examples and challenges, many of which are concerned with the handling of the so-called contrary-to-duty (CTD) reasoning and deontic conflicts. A whole plethora of formal methods have been developed to handle CTD and deontic conflicts, methods that go far beyond simple modal logics such as SDL (standard deontic logic). Furthermore, it is widely held that norms are defeasible and come with exceptions and priorities. The seminar participants discussed the role of nonmonotonicity in deontic logic and the use of techniques from formal argumentation to define *defeasible* deontic logics.

From theory to tools. The fourth topic of the seminar concerned experimenting and implementing normative reasoning. One of the themes discussed had to do with integrating normative reasoning techniques with reinforcement learning (RL) in the design of ethical autonomous agents. Another theme that was discussed had to do with the automatization of deontic explanations. For example, in the recently introduced Logikey framework, it has been shown how Isabelle/HOL can be used as flexible interactive testbed for the design of domain-specific logical formalisms. Isabelle/HOL incorporates a number of automated tools that provide just-in-time feedback (counter-models, examples, proofs) to the formalization process. This feedback can be used to assess and reflect upon the theoretical properties of the system being designed/implemented. We can encode complex semantics in Isabelle/HOL as well as notions of argumentation (already partly done for abstract argumentation) so that Isabelle/HOL is turned into a reasoning system for those specific formalisms. What's more; notions of deontic explanations can be encoded and experimented with. Another key tool for automatize normative reasoning is analytic proof systems, which were also discussed in the seminar.

2 Table of Contents

3 Overview of Talks

3.1 Principles for a judgement editor based on Binary Decision Diagrams

Guillaume Aucher (University of Rennes, FR)

Joint work of Guillaume Aucher, Anthony Baire, Jean Berbinau, Annie Foret, Jean-Baptiste Lenhof, Marie-Laure Morin, Olivier Ridoux, François Schwarzentruber
Main reference Guillaume Aucher, Jean Berbinau, Marie-Laure Morin: "Principles for a Judgement Editor Based on Binary Decision Diagrams", Journal of Applied Logics -IfCoLog Journal of Logics and their Applications, Vol. 6(5), p. 33, 2019.
URL https://inria.hal.science/hal-02273483

We introduce the theoretical principles that underlie the design of a software tool which could be used by judges for making decisions about litigations and for writing judgements. The tool is based on Binary Decision Diagrams (BDD), which are graphical representations of truth-valued functions associated to propositional formulas. Given a type of litigation, the tool asks questions to the judge; each question is represented by a propositional atom. Their answers, true or false, allow to evaluate the truth value of the formula which encodes the overall recommendation of the software about the litigation. Our approach combines some sort of "theoretical" or "legal" reasoning dealing with the core of the litigation itself together with some sort of 'procedural' reasoning dealing with the protocol that has to be followed by the judge during the trial: some questions must necessarily be examined and sometimes in a specific order. That is why we consider extensions of BDD called Multi-BDD. They are BDD with multiple roots corresponding to the different specific issues that must necessarily be addressed by the judge during the trial. We illustrate our ideas on a case study dealing with French trade union elections which has been used throughout our project with the Cour de cassation. We also introduce the prototype developed during our project and a link with restricted access to try it out.

3.2 The moral disconnect in LLMs

Jan M. Broersen (Utrecht University, NL)

I will point out what is wrong with the moral behavior of LLMs like ChatGPT. Then I will ponder the question if we can actually solve the moral disconnect observed.

Large language model-based artificial conversational agents (like ChatGPT) can answer ethical questions. Just on the basis of that capacity, we may attribute a weak form of ethical knowledge to them. But do these models use this knowledge as a basis for their own ethical behaviour? I argue that cannot be the case. I will refer to this failure as the "ethical knowledge disconnect" of LLM-based agents. To understand the disconnect, we have to understand how ethical behavioural "guardrails" are implemented in systems like ChatGPT. I argue all methods currently employed do little to solve the disconnect. I wil also discuss how the disconnect may extend to non-ethical behaviours and should rather be seen as an instance of a more general knowledge disconnect. If that is the case, there are implications for making LLMs the basis for embodied agents. Finally, I will report on my attempt to expose the disconnect by trying to force ChatGPT into an ethical performative contradiction.

3.3 Machine ethics and precedent-based reasoning

Ilaria Canavotto (University of Maryland – College Park, US)

Joint work of Ilaria Canavotto, John Horty, Eric Pacuit
Main reference Ilaria Canavotto, John F. Horty: "Piecemeal Knowledge Acquisition for Computational Normative Reasoning", in Proc. of the AIES '22: AAAI/ACM Conference on AI, Ethics, and Society, Oxford, United Kingdom, May 19 – 21, 2021, pp. 171–180, ACM, 2022.
URL https://doi.org//10.1145/3514094.3534182

I will present research that I am carrying out in collaboration with John Horty and Eric Pacuit. We are exploring a hybrid approach to knowledge acquisition and representation for computational normative reasoning (a.k.a. machine ethics). Building on recent research in artificial intelligence and law, our approach is modeled on the familiar practice of decision-making under precedential constraint in the common law. I will first introduce a formal model of this practice (called the reason model of precedential constraint), showing how a body of normative information can be constructed in a way that is piecemeal, distributed, and responsive to particular circumstances. I will then discuss a possible application to the design of a robot childminder.

3.4 Data-driven norm revision

Mehdi Dastani (Utrecht University, NL)

Joint work of Davide Dell'Anna, Natasha Alechina, Fabiano Dalpiaz, Mehdi Dastani, Brian Logan
Main reference Davide Dell'Anna, Natasha Alechina, Fabiano Dalpiaz, Mehdi Dastani, Brian Logan: "Data-Driven Revision of Conditional Norms in Multi-Agent Systems", J. Artif. Intell. Res., Vol. 75, pp. 1549–1593, 2022.
URL https://doi.org//10.1613/jair.1.13683

Norm enforcement is a mechanism for steering the behavior of individual agents to achieve desired system-level objectives. Due to the dynamics of systems, however, it is hard to design norms that guarantee the achievement of the objectives in every operating context. In this work, we propose a data-driven approach to norm revision that synthesises revised norms with respect to a data set consisting of traces describing the behavior of the individual agents in the system. The proposed approach synthesises revised norms that are significantly more accurate than the original norms in distinguishing adequate and inadequate behaviors for the achievement of the system-level objectives.

3.5 Normative reasoning and the UK Highway Code

Louise A. Dennis (University of Manchester, GB)

Joint work of Joe Collenette, Louise A. Dennis, Michael Fisher
Main reference Joe Collenette, Louise A. Dennis, Michael Fisher: "Advising Autonomous Cars about the Rules of the Road", in Proc. of the Proceedings Fourth International Workshop on Formal Methods for Autonomous Systems (FMAS) and Fourth International Workshop on Automated and verifiable Software sYstem DEvelopment (ASYDE), FMAS/ASYDE@SEFM 2022, and Fourth International Workshop on Automated and verifiable Software sYstem DEvelopment (ASYDE)Berlin, Germany, 26th and 27th of September 2022, EPTCS, Vol. 371, pp. 62–76, 2022.
URL https://doi.org//10.4204/EPTCS.371.5

Our recent formalisation of the UK Highway Code has highlighted a number of ways normative reasoning interacts with it. Of particular interest are rules that implicitly defer to norms: e.g, "be considerate to other road users", many rules that have normative rather than legal force (the code distinguishes between rules that legally "must" be obeyed and rules that normatively "should" be obeyed), as well as rules which allow things which would be normatively impermissible (driving at night with the headlights off in well-lit areas). This opens up a defined area of computer reasoning in which the interaction of legal and normative rules can be studied.

3.6 Witnesses and explanations for answer set programming

Thomas Eiter (TU Wien, AT)

Joint work of Yisong Wang, Thomas Eiter, Yuanlin Zhang, Fangzhen Lin
Main reference Yisong Wang, Thomas Eiter, Yuanlin Zhang, Fangzhen Lin: "Witnesses for Answer Sets of Logic Programs", ACM Trans. Comput. Log., Vol. 24(2), pp. 15:1–15:46, 2023.
URL https://doi.org//10.1145/3568955

Answer Set Programming (ASP) is a popular declarative problem solving paradigm that has been widely applied in various domains. Given that answer sets are supposed to yield solutions to the original problem, the question of "why a set of atoms is an answer set" becomes important for both semantics understanding and program debugging. In this talk, we briefly consider recent work on answering such questions on disjunctive logic programs, as a basis for building explanations on top of ASP programs.

References

1 Yisong Wang, T. Eiter, Y. Zhang, and F. Lin. Witnesses for answer sets of logic programs. *ACM Transactions on Computational Logic*, 24(2), Apr. 2023. Article no. 15, 46 pp.

3.7 Machine learning with (logical) requirements

Eleonora Giunchiglia (TU Wien, AT)

Joint work of Eleonora Giunchiglia, Fergus Imrie, Mihaela van der Schaar, Thomas Lukasiewicz, Mihaela Stoian, Salman Khan, Fabio Cuzzolin

Machine learning models have revolutionised various fields by providing highly effective solutions to complex problems. However, their success comes at the cost of unexpected behaviours, which might violate known requirements expressing background knowledge about the problem at hand. This can have dramatic consequences, especially in safety critical scenarios (e.g., healthcare/autonomous driving). In this talk, I will first give an overview of the standard performance-driven machine learning development pipeline, and then I will present our proposed requirements-driven machine learning development process, highlighting its advantages. Finally, I will argue that it is desirable to use logic to express requirements, and I will briefly discuss how different neuro-symbolic methods have been developed to incorporate both norms (or soft constraints) and requirements (or hard constraints).

References

1 Eleonora Giunchiglia, Fergus Imrie, Mihaela van der Schaar, Thomas Lukasiewicz. *Machine Learning with Requirements: a Manifesto.* https://arxiv.org/pdf/2210.01597.pdf, 2023
2 Eleonora Giunchiglia, Mihaela Cătălina Stoian, Salman Khan, Fabio Cuzzolin, Thomas Lukasiewicz. *ROAD-R: The Autonomous Driving Dataset with Logical Requirements.* Machine Learning Journal, 2023
3 Eleonora Giunchiglia, Thomas Lukasiewicz. *Multi-Label Classification Neural Networks with Hard Logical Constraints.* Journal of Artificial Intelligence Research, 2021
4 Eleonora Giunchiglia, Mihaela Cătălina Stoian, Thomas Lukasiewicz. *Deep Learning with Logical Constraints.* IJCAI, 2022

3.8 Is the Chisholm paradox a paradox?

Guido Governatori (Tarragindi, AU)

Main reference Guido Governatori: "A Short Note on the Chisholm Paradox", in Proc. of the 4th International Workshop on MIning and REasoning with Legal texts co-located with the 32nd International Conference on Legal Knowledge and Information Systems (JURIX 2019), Madrid, Spain, December 11, 2019, CEUR Workshop Proceedings, Vol. 2632, CEUR-WS.org, 2019.
URL https://ceur-ws.org/Vol-2632/MIREL-19_paper_4.pdf

We advance an alternative version of the Chisholm Paradox and we argue that the alternative version (while logically equivalent to the original version), in its manifestation in the natural language, is not intuitively consistent. The alternative version of the paradox suggests some requirements for deontic logics designed for legal reasoning.

References

1 José Carmo and Andrew J. I. Jones. *Deontic Logic and Contrary-to-Duties*, Handbook of Philosophycal Logic (2nd edition), Volume , pages 265–343. Springer Netherlands, Dordrecht, 2002.
2 Rodrick M. Chisholm. Contrary-to-Duty Imperatives and Deontic Logic. *Analysis*, 24(2):33–36, 1963.

3 Guido Governatori. A short note on the Chisholm Paradox. Proceedings of the 4th International Workshop on MIning and REasoning with Legal texts, CEUR-Workshop Proceedings 2632.
4 John Horty. Deontic modals: Why abandon the classical semantics. *Pacific Philosophical Quarterly*, 95:424–460, 2014.
5 James E. Tomberlin. Contrary-to-duty imperatives and conditional obligation. *Noûs*, 15(3):357–375, 1981.
6 Lennart åqvist. Good Samaritan, contrary-to-duty imperatives, and epistemic obligations. *Noûs*, 1(4):361–379, 1967.

3.9 n problems for deontic logic for normative reasoning.

Guido Governatori (Tarragindi, AU)

Main reference Guido Governatori: "Thou shalt is not you will", in Proc. of the 15th International Conference on Artificial Intelligence and Law, ICAIL 2015, San Diego, CA, USA, June 8-12, 2015, pp. 63–68, ACM, 2015.
URL https://doi.org//10.1145/2746090.2746105

The original intent of this talk was to highlight some problems/issues a deontic logic has to address to capture normative (legal) reasoning (including some that might be controversial). However, after some of the previous presentation, the focus shifted to one of the issues, more specifically whether it is possible to use other logic, in particular Temporal Logic to model legal reasoning. I show that using Temporal Logic, more precisely, Linear Temporal Logic, as done by some work in the area of business process compliance, leads to some paradoxical results: either it is not possible to model some deontic aspects, or the outcome of the modelling contradicts expected legal outcome.

References

1 Guido Governatori. Thou Shalt is not You Will. Proceedings International Conference on Artificial Intelligence and Law 2015, pp. 63-68. doi: 10.1145/2746090.2746105
2 Guido Governatori and Mustafa Hashmi. No Time for Compliance. Proceedings EDOC 2015: pp. 9-18 doi: 10.1109/EDOC.2015.12

3.10 Deontic explanation: questions, dilemma's and choice

Joris Hulstijn (University of Luxembourg, LU) and Leendert van der Torre (University of Luxembourg, LU)

When a computer system takes decisions that affect people, they may demand an explanation. When the application involves norms, we need a deontic explanation. An deontic explanation is analyzed here as an answer to a why-question, relative to a normative system. Just like a who-question asks for persons, a why-question asks for reasons. We analyze differences and similarities of three kinds of semantics, that are formulated in terms of a partition of the set of possible worlds: (1) questions and answers, (2) moral dilemmas, and (3) see-to-it-that choice structures. The analysis is built on an analogy between providing an answer to a question, resolving a moral dilemma and choosing an action. The role of the context in these types of semantics can be naturally analysed in a form of update semantics. In future work we hope to find constructions in the object language, to construct or reframe a question, a choice for action, or a dilemma, and ways of answering or resolving them.

3.11 The logic of second-order reasons

Aleks Knoks (University of Luxembourg, LU)

A normative reason is a consideration that counts either in favor of or against an action or attitude. A second-order normative reason, then, is a consideration that counts in favor of or against taking another consideration to be a normative reason. While some authors have questioned the existence of such reasons, others assign them very important roles. Thus, exclusionary or negative second-order reasons – that is, reasons against taking other considerations to be reasons – play a crucial role in Joseph Raz's account of practical reasoning. The primary goal of this talk is to show how second-order reasons and their normative effects can be captured in default logic. Starting with Horty s default logic-based model of the way reasons interact to support ought statements, I explain why one can't rest content with Horty's formalization of exclusionary reasons. Most importantly, it assimilates defeat by exclusionary reasons to canceling (or undercutting), doesn't do justice to the idea that excluded first-order reasons remain valid, and doesn't account for a distinct sense of "ought" grounded in first-order reasons. I discuss an alternative model, present an account of positive-second order reasons, and explore the model's predictions regarding the structure of even higher-order reasons and conflicts between them.

3.12 Moral planning agents

Emiliano Lorini (CNRS – Toulouse, FR)

The talk shows how non-classical logics with special emphasis on modal logic, epistemic logic and conditional logic can be used to represent and compare a rich variety of explanations of classifier systems; these include abductive, contrastive, counterfactual, objective vs subjective, and interactive explanations. The first part of the presentation will be devoted to explaining "white box" classifiers that are assumed to be perfectly known, while the second part will focus on "black box" classifiers about which the external observer has only partial knowledge. I will present proof-theoretic and complexity results for the involved logics and illustrate their expressiveness through concrete examples.

References

1 Liu, X., Lorini, E. (2023). A Unified Logical Framework for Explanations in Classifier Systems. Journal of Logic and Computation, 33(2), pp. 485-515
2 Liu, X., Lorini, E. (2022). A Logic of "Black Box" Classifier Systems. In Proceedings of the 28th Workshop on Logic, Language, Information and Computation (WOLLIC 2022), LNCS, volume 13468, Springer-Verlag, pp. 158–174

3 Liu, X., Lorini, E. (2021). A Logic for Binary Classifiers and Their Explanation. In Proceedings of the 4th International Conference on Logic and Argumentation (CLAR 2021), LNCS, volume 13040, Springer-Verlag, pp. 302-321.
4 Aguilera-Ventura, C., Herzig, A., Liu, X., Lorini, E. (2023). Counterfactual Reaasoning via Grounded Distance. In Proceedings of the 20th International Conference on Principles of Knowledge Representation and Reasoning (KR 2023), forthcoming.

3.13 How to implement cognitive and social properties of norms in robots: The promise of behavior trees

Bertram F. Malle (Brown University – Providence, US)

Joint work of Bertram F. Malle, Eric Rosen, Vivienne B. Chi, Dev Ramesh
Main reference Bertram F. Malle, Eric Rosen, Vivienne B. Chi, Dev Ramesh: "What properties of norms can we implement in robots¿' Proceedings of the 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN 2023), August 2023.

Norms are indispensable for human communities, and so they will be for robot-human communities. We analyze some of the requirements for a robot to have norms and conform its actions to them. These requirements include both cognitive and social properties that human norms have. We examine which of these properties can be implemented in a robot's architecture and review some previous computational approaches. We then introduce a new one using behavior trees, argue for its promise to implement properties of norms, and discuss unsolved challenges.

3.14 Towards a mechanisation of the proof theory of normative reasoning

Xavier Parent (TU Wien, AT)

Joint work of Xavier Parent, Agata Ciabattoni, Nicola Olivetti
Main reference Agata Ciabattoni, Nicola Olivetti, Xavier Parent: "Dyadic Obligations: Proofs and Countermodels via Hypersequents", in Proc. of the PRIMA 2022: Principles and Practice of Multi-Agent Systems – 24th International Conference, Valencia, Spain, November 16-18, 2022, Proceedings, Lecture Notes in Computer Science, Vol. 13753, pp. 54–71, Springer, 2022.
URL https://doi.org//10.1007/978-3-031-21203-1_4

This work lays the groundwork for a (proper) mechanisation of normative reasoning, via the use of a so-called analytic sequent calculi. They are particularly useful for backward reasoning and deontic explanations. To answer a question of the form "why should I do X?", needed is to retrieve the path (the "proof") leading to this conclusion. Analytic calculi allow precisely this.

This is part of a bigger project aiming at developing analytic proof systems for deontic logic formalisms.

We consider a SoA formalism, the preference-based system E for conditional obligation due to Aqvist. Its key strength lies in its ability to resolve the CTD paradox. We provide an analytic calculus for it, the first of its kind. We also provide a terminating countermodel generation procedure in case of failure of proof search, and a complexity result (co-NP).

3.15 Deontic to description logics

Bijan Parsia (University of Manchester, GB)

Joint work of Bijan Parsia, E'leanor Turner, Ui Sattler

Deontic logic is a family of often propostional modal logics intended to capture certain kinds of moral reasoning, centrally those involving obligation. Deonitc logic has receive a great deal of attention from philosophical logicians and there has been some interest in developing implementations of reasoning procedures for various deontic logics. However, much of the work has rested on axiomatic approaches with the inference rules of necessitation and *modus pones*. While familiar and convenient for some communities, these not a standard basis for robust implementations.

Description logics are a widespread family of decidable logics, the core of which can be seen as notational variants of propositional modal logics. The description logic $\mathcal{SHROIQ}$ forms the logical foundation of the standardised ontology language OWL 2 DL. OWL 2 has broad and deep infrastructure including production quality reasoners, IDEs, services, repostiories, and so on.

Give the very expressivity and wide range of available tools, description logics are an attractive foundation for a translational approach to implementing reasoning services for deontic logics.

We present various translations of several logics of obligation and agency into OWL 2. While all the translations preserve the meaning (at least in the sense of always being able to recover all entailments) of the original, they have different characteristics which affect the performance of automated reasoning tasks and the utility of the results for users.

We also explore the benefits of an ontology-oriented approach to modeling deontic problems.

3.16 KI Wissen – Development of methods for integrating knowledge into machine learning in autonomous driving

Adrian Paschke (FU Berlin, DE)

Main reference Julian Wörmann, Daniel Bogdoll, Etienne Bührle, Han Chen, Evaristus Fuh Chuo, Kostadin Cvejoski, Ludger van Elst, Tobias Gleißner, Philip Gottschall, Stefan Griesche, Christian Hellert, Christian Hesels, Sebastian Houben, Tim Joseph, Niklas Keil, Johann Kelsch, Hendrik Königshof, Erwin Kraft, Leonie Kreuser, Kevin Krone, Tobias Latka, Denny Mattern, Stefan Matthes, Mohsin Munir, Moritz Nekolla, Adrian Paschke, Maximilian Alexander Pintz, Tianming Qiu, Faraz Qureishi, Syed Tahseen Raza Rizvi, Jörg Reichardt, Laura von Rüden, Stefan Rudolph, Alexander Sagel, Gerhard Schunk, Hao Shen, Hendrik Stapelbroek, Vera Stehr, Gurucharan Srinivas, Anh Tuan Tran, Abhishek Vivekanandan, Ya Wang, Florian Wasserrab, Tino Werner, Christian Wirth, Stefan Zwicklbauer: "Knowledge Augmented Machine Learning with Applications in Autonomous Driving: A Survey", CoRR, Vol. abs/2205.04712, 2022.
URL https://doi.org//10.48550/arXiv.2205.04712

AI-based processes are paving the way to fully automated autonomous driving. Up until now, the development of AI solutions has been purely driven by data. This data driven approach requires enormous amounts of data for the training and validation of AI functions, with the collection and processing of this data being very resource-intensive and expensive. In addition to the dependence on extensive amounts of data, data-based AI processes have another weakness: they are still generally black-box models for which the decision making process cannot be directly reconstructed. In the talk I will report about the project "KI Wissen" (`https://www.kiwissen.de/`) and neuro-symbolic methods for integrating existing knowledge into the data-driven AI functions of autonomous vehicles (AVs) and vice versa for extracting interpretable symbolic knowledge from deep neural network models. In this talk I will specifically present an approach for extracting interpretable hierarchical rules from the learned AVs' deep neural networks and a neuro-symbolic architecture for a hybrid integration of deep neural networks, modelling the AVs' behaviour and situation information, with symbolic knowledge models for representing ontological domain and world knowledge for situation interpretation and for representing rule-based legal knowledge and norms for legal reasoning and compliance checks. The goal of the KI Wissen project is to create a comprehensive ecosystem for the integration of knowledge into the training and safeguarding of AI functions. By combining conventional data-based AI methods with the knowledge- or rule-based methods developed in the project, the basis for training and validating of AI functions will be completely redefined: This basis now includes not only data, but information, i.e., data and knowledge. The development from data- to information-based AI carried out in the project addresses the central challenges towards autonomous driving: the generalization of AI to phenomena with small data bases, the increase of the stability of the trained AI to disturbances in the data, the data efficiency, the plausibility check and the validation of AI-supported functions as well as the increase of the functional quality.

3.17 Legal explanations

Antonino Rotolo (University of Bologna, IT)

One fundamental question lies behind the distinction between justification and explanation in normative reasoning. In fact, we must notice that the tradition of legal logic and legal theory, in modeling legal decision-making, very often elaborate on various types of justification and takes this last concept as central, somehow maintaining that the idea of explanation depends on justification: while the explanation of a legal decision does not necessarily correspond to a justificatory reason for it, the opposite usually holds. The talk will offer some formal insights about this topic.

3.18 Combining Deep NLP with symbolic reasoning in automatic legal judgement

Ken Satoh (National Institute of Informatics – Tokyo, JP)

Joint work of Ha-Thanh Nguyen, Wachara Fungwacharakorn, Fumihito Nishino, Ken Satoh, Magumi Fujita
Main reference Ha-Thanh Nguyen, Wachara Fungwacharakorn, Fumihito Nishino, Ken Satoh: "A Multi-Step Approach in Translating Natural Language into Logical Formula", in Proc. of the Legal Knowledge and Information Systems – JURIX 2022: The Thirty-fifth Annual Conference, Saarbrücken, Germany, 14-16 December 2022, Frontiers in Artificial Intelligence and Applications, Vol. 362, pp. 103–112, IOS Press, 2022.
URL https://doi.org//10.3233/FAIA220453

We show how to combine deep NLP with nonmonotic reasoning in legal domain. We extract legally relevant facts from a case description written in natural language using deep NLP and input these facts into manually encoded articles in our legal logic programming language PROLEG to make legal judgement and produce explanation of the judgement.

3.19 What are social norms?

Kai Spiekermann (London School of Economics, GB)

Main reference Kai Spiekermann: "Review: Explaining Norms, Geoffrey Brennan, Lina Eriksson, Robert E. Goodin and Nicholas Southwood", Oxford University Press, 2013, Economics and Philosophy, vol. 31, no. 1.
URL https://www.kaispiekermann.net/blog-native/2015/6/9/book-review-explaining-norms-geoffrey-brennan-lina-eriksson-robert-e-goodin-and-nicholas-southwood

What are social norms? There is a surprising level of disagreement about this question in the literature. I compare Bicchieri's (2003) game-theoretic account with Brennan, Erikson, Goodin, and Southwood's account of norms as cluster of attitudes. Both accounts agree that real or perceived social practices and normative expectations play a central role. However, examples show that the different accounts can come apart.

4 Working groups

4.1 Explanation in case-based reasoning

Ilaria Canavotto (University of Maryland – College Park, US), John F. Horty (University of Maryland – College Park, US), Bijan Parsia (University of Manchester, GB), and Henry Prakken (Utrecht University, NL)

Computational models of legal precedent-based reasoning developed in the field of Artificial Intelligence and Law have recently been applied to the development of explainable AI methods. The key idea behind this approach is to interpret training data as a set of precedent cases; a model of precedent-based reasoning can then be used to build either an interpretable system for binary classification [1, 2] or an algorithm that generates post-hoc justifications for the decisions of a machine learning system for binary classification [3]. This breakout session has been devoted to discuss a number of technical and conceptual questions concerning the framework for post-hoc justification proposed in [3].

References

1 Cocarascu, O. Čyras, K. and Toni, F. *Explanatory predictions with artificial neural networks and argumentation*. IJCAI/ECAI-2018 Workshop on Explainable Artificial Intelligence, pp. 26-32.
2 Čyras, K., Satoh, K. and Toni, F. *Explanation for case-based reasoning via abstract argumentation*. COMMA 2016, pp. 26–32.
3 Prakken, H. and Ratsma, R. *A top-level model of case-based argumentation for explanation: Formalisation and experiments*, Argument & Computation 13, 2022, pp. 159–194.

4.2 Justification and explanation

Ilaria Canavotto (University of Maryland – College Park, US), Pedro Cabalar (University of Coruña, ES), Thomas Eiter (TU Wien, AT), Joris Hulstijn (University of Luxembourg, LU), Aleks Knoks (University of Luxembourg, LU), Eric Pacuit (University of Maryland – College Park, US), Bijan Parsia (University of Manchester, GB), Henry Prakken (Utrecht University, NL), and Antonino Rotolo (University of Bologna, IT)

The problem of developing explainable AI methods is becoming increasingly central in AI. At the same time, the notions of explanation, explainability, and justification have been extensively investigated in philosophy, law, and social science. As a result, an increasing number of scholars from these disciplines is considering how to apply research in these fields to explainable AI. One problem of this interdisciplinary effort is that, more often than not, researchers from different fields have different understandings of what the task underlying explainable AI is (or should be) or what exactly "explanation" or "justification" mean when applied to AI systems. This working group aimed at identifying some key distinctions that could be used to build a unified conceptual framework for explainable AI. The discussion was split into two parts:

Part 1: Initial definitions. Most participants agreed that, when discussing explainable AI methods, it is helpful to introduce an initial distinction between explanation, explication, and justification. Although there was substantial disagreement about the exact definition of each notion, we agreed on the following prelimary characterization:

Explanation is about how a system reached a particular decision given a certain input and what the reason (motive, or cause) of the decision was. Explanation is important for bias detection.

Explication aims at making the user understand how the system behaves.

Justification is about finding an argument why a particular decision is reasonable or normatively acceptable. Justifications are post-hoc.

Part 2: Tasks underlying explainable AI. Some participants suggested that, in order to build a unified conceptual framework for explainable AI, distinguishing the different tasks that fall under the label "explanation in AI" might be more effective than finding satisfactory definitions of the notions above. We discussed this issue by taking, as a toy example, a system that takes as input a patient's description and returns as output a prediction of whether the patient will be alive in five years. The tasks we identified are as follows:

1. Extract the mechanism that leads from input to output. This task only applies to the case in which the system we are working with is interpretable. The aim underlying the task is to understand how the system produces a prediction. The target (or audience) are the designers of the system. Importantly, the specific form and level of abstraction of the extracted mechanism depends on the designer and what specifically they want to understand. For instance, suppose that the system is based on case-based reasoning but, because of their training, the designer understands abstract argumentation theory better than case-based reasoning. Then the designer might extract the mechanism underlying the system by mapping the case-based reasoning system into abstract argumentation theory.
2. Make the system "user friendly." Once the designer has extracted the mechanism underlying the system, there is a further question of how to make this mechanism accessible to a user. Continuing on the example of a case-based reasoning system, a designer might understand how the system works by proving that the system reaches a decision when, say, the grounded extension of the argumentation framework the system was mapped to contains certain arguments. Of course, things are different for the user, who has probably never heard of abstract argumentation frameworks and grounded semantics. But the designer could make the system "user friendly" by making it capable, first, of extracting an argumentative explanation from the abstract argumentation framework in question and, second, of producing a text containing a translation of the explanation in natural language. As before, the specific form and level of abstraction of the generated explanation depends on the target user and the context.
3. Compare the extracted mechanism with other mechanisms. While tasks 1 and 2 are about understanding how the system works, this task is about finding reasons to accept the predictions of the system. In the toy example of a system that predicts whether a patient will be alive in five years, the task could be understood as comparing the answers to the questions "why did the system predicted that the patient will be alive in five years?" and "why is it reasonable to think that the patient will be alive in five years?". Comparing the answers to the two questions is a way to assess the trustworthiness of the system. In case the system we are working with is a black box and it is not possible to answer the first question, answering the second question is a way to produce a post-hoc justification of the predictions of the system.
4. Extract normative justifications. In the normative domain, justified decisions are decisions that comply with a set of norms. In case the system we are working with is trained on a normative dataset or the mechanism underlying it is subject to normative constraints, then an additional task is to justify the system's decisions by verifing that they were reached without violating the relevant norms.

4.3 Modeling normative reasons

Aleks Knoks (University of Luxembourg, LU), Christoph Benzmüller (Universität Bamberg, DE), Huimin Dong (Sun Yat-Sen University – Zhuhai, CN), Joris Hulstijn (University of Luxembourg, LU), Eric Pacuit (University of Maryland – College Park, US), Antonino Rotolo (University of Bologna, IT), Christian Straßer (Ruhr-Universität Bochum, DE), and Leendert van der Torre (University of Luxembourg, LU)

When philosophers talk about normative matters – about what is right, obligatory, permitted, and so on – they tend to rely on the notion of a normative reason. In the practical domain – which includes morality, as well as the domain of practical rationality – normative reasons are understood as considerations that count in favor of or against actions. The notion has become a mainstay of philosophy, where it is very often relied on in answering various normative and metanormative questions. The so-called reasons-first program takes this to the extreme, taking the notion of a normative reason to be basic and holding that all other normative notions are to be analyzed in terms of it [1, 2, 3]. When discussing the interaction between reasons, the philosophical literature uses such phrases as "the action supported on the balance of reasons" and "reasons for outweigh reasons against", inviting an image of a weighing scale. Philosophers have explored various ideas about the exact workings of this normative weighing scale, with rare exceptions, their investigations have been carried out informally. The overall goal of this breakout session, then, was to model normative reasons and their interaction – roughly, the normative weighing scale for reasons – using methods from mathematical modeling and knowledge representation.

Toward the end of the discussion, the following model emerged:

A structure $T = \langle P, I, F, f, N, D\rangle$ is a model of reasons, where:

- P is a set of persons;
- I is a set of issues;
- F is a set of features;
- $f : P \times I \to 2^F$ is a function mapping pairs of persons and issues to a set of features, indicating which features serve as normative reasons in determining the normative status of the issue for the given person;
- N is a set of polarity functions, with each element n of N having the form $n : P \times I \times \{f\} \times F \to \{+, -, 0\}$ (as their name suggests, these functions determine the polarity or directedness of features);
- D is a set of deontic functions, with each element d of D having the form $d : P \times I \times f \times N \to \{+, -, 0\}$ (as the name suggests, these functions determine the normative status of issues for persons: intuitively, an assignment of $+$ means that the issue (action) is obligatory for the person, that of $-$ means that it is forbidden, and that of 0 means that it is indifferent).

It was noted that, its simplifying assumptions notwithstanding, this model captures important parts of philosophers' way of thinking about normative reasons, the interaction between reasons (or weighing reasons), and the relation between reasons and such normative notions as obligations and permissions. It was also noted that this model is only a starting point, and that more structure can be added to it with ease. For instance, one could make deontic functions depend on additional arguments (representing other normatively relevant information), or one could allow for multiple types of deontic functions (representing different types of obligations). One could also substitute the polarity functions with numerical functions, with the result that the features identified as reasons for (against) an issue would be associated not only with polarities, but also with magnitudes, bringing the model even closer to the metaphor of weight scales. The affinities with the field of multi-criteria decision-making [4] were also noted.

The discussion participants agreed that the model can be used to formalize various sorts of methodological questions, making them (more) tractable. Questions prompted by the talks delivered at the seminar served as examples: What is the role of polarity (or "directedness") in reason-based decisions? Are Raz's views on reasons in the practical domain equivalent to Pollock's views in the epistemic domain?

For completeness, it should be added that the breakout session participants also voiced some reservations toward the model. Thus, it was noted that the model does not (yet) represent normatively relevant considerations that are not reasons – including conditions and modifiers – which are widely discussed in the philosophical literature. Another issue that was noted was that the model allows one to represent only situations of binary choice. Still, all participants agreed that these issues can be overcome.

References

1. Scanlon T. M. *What We Owe to Each Other*. Harvard University Press, 1998.
2. Raz J. *Practical Reason and Norms*. Oxford University Press, 1990.
3. Parfit D. *On What Matters*, vol. I, Oxford University Press, 2011.
4. Keeney R. and Raiffa H. *Decisions with Multiple Objectives*. Cambridge University Press, 1993.

4.4 Normative reasoning for autonomous agents (Parts I and II)

Pedro Cabalar (University of Coruña, ES), Agata Ciabattoni (Vienna University of Technology, AT), Mehdi Dastani (Utrecht University, NL), Louise A. Dennis (University of Manchester, GB), Huimin Dong (Sun Yat-Sen University – Zhuhai, CN), Thomas Eiter (Vienna University of Technology, AT), Eleonora Giunchiglia (Vienna University of Technology, AT), Guido Governatori (Tarragindi, AU)

From self-driving cars and unmanned aerial vehicles to robot nannies and elder care robots, the myriads of practical uses of Artificial Intelligence (AI) only continue to grow. In these applications an increasingly prominent role is played by autonomous agents, which should operate in an "intelligent" way on some users's behalf but without human intervention. Autonomous agents must accomplish a variety of real world tasks and need to adapt to potentially unpredictable changes in their environment. Reinforcement Learning (RL) – a prominent machine learning technique – has demonstrated to be an effective tool for teaching agents such behaviour [1].

As we assign more roles to RL-based agents it becomes crucial to ensure that they act in ways that are legal, ethics-sensitive, and socially acceptable. This introduces a further challenge: establishing boundaries around the behaviour of these agents, i.e. equipping them with the ability to comply with legal, ethical and social norms, while still enacting pre-learned optimal behaviour.

We have recognized the different approaches that emerge and did thoroughly discuss them. The first approach uses symbolic AI techniques (a.k.a. Logic, Knowledge Representation and Reasoning) and was successfully employed, e.g., in [2] where a theorem proved for a defeasible deontic logic advises the learning agent on the compliant actions; this approach can however be computationally expensive and less suited for dealing with (signal-based) data, sensory input, or stochastic environments. The other approach relies on sub-symbolic AI (a.k.a. Machine Learning), and was applied, e.g., in [3], to constraint the behaviour of AI agents via reward/penalities; this approach excels under these conditions and enables the construction of efficient and adaptable AI systems, which however lack modularity and transparency; moreover, it is not clear how to adapt this approach to deal with complex normative systems.

All participants agreed that the best way to proceed would be to interlace the two approaches, thus providing the best of both worlds. This breakout session consisted of two parts.

PART I. The participants have identified the main steps for achieving this very challenging task: first translate the norms into efficiently computable representations of normative knowledge, and afterwards to exert some form of normative reasoning to be integrated with the agent's training. Concrete candidates for the norm translations have been proposed: Answer Set Programming [4] and computationally-oriented deontic logics (e.g., Defeasible Deontic Logic [6]).

PART II. (A subgroup[1] of the) participants have concretely discussed potentially useful tools for the second step, and also feasible case studies that could be employed to test their effectiveness and feasibility. There was consensus that the integration of normative reasoning and the Machine Learning component would be the most complex part of the enterprise. To this aim the participants agree that it is worth trying to adapt/extend the techniques used in the Safe Reinforcement Learning community, and/or the emerging idea of constraining Machine Learning with logical formulas [5].

References

1. D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. P. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, and D. Hassabis. Mastering the game of Go without human knowledge. Nat., 550(7676):354–359, 2017.
2. E. A. Neufeld, E. Bartocci, A. Ciabattoni, and G. Governatori. Enforcing ethical goals over reinforcement-learning policies. J. of Ethics and Inform. Techn., 2022
3. R. Noothigattu, D. Bouneffouf, N. Mattei, R. Chandra, P. Madan, K. R. Varshney, M. Campbell, M. Singh, and F. Rossi. Teaching AI agents ethical values using reinforcement learning and policy orchestration. In Proc. IJCAI, 2019.
4. G. Brewka, T. Eiter, and M. Truszczynski, editors. AI Magazine: special issue on Answer Set Programming. AAAI Press, 2016. Volume 37, number 3. Editorial pp. 5-6.
5. E. Giunchiglia, M. Stoian, T. Lukasiewicz: Deep Learning with Logical Constraints. IJCAI 2022: 5478-5485
6. G. Governatori: Practical Normative Reasoning with Defeasible Deontic Logic. Reasoning Web 2018: 1-25

[1] Some of the participants had moved to other breakout sessions.

4.5 The normative competence of artificial agents

Kevin Baum (University of Saarbrucken, DE), Jan Broersen (Utrecht University, N), Louise A. Dennis (University of Manchester, GB), Frank Dignum (Umeå University, SE), Virginia Dignum (Umeå University, SE), Bertram Malle (Brown University, USA), Xavier Parent (Vienna University of Technology, A), Marija Slavkovik (University of Bergen, NO), Kai Spiekermann (London School of Economics, UK)

This scenario seems to repeat itself: A company creates a service that uses artificial intelligence and/or has some agency. We will refer to this service very generically as “a machine”. In its interaction with the users, this machine inevitably violates a norm. The company responds with a constraints that disables the machine from violating the norm, effectively turning the norm into a constraint. Then a new situation arises in which the machines constrained behaviour violates another norm. This practice does not result with a machines ability to operate in a normative context, but rather with a system whose behaviour is neither desirable, nor predictable. To be able to move away from this trap of update and adjust, we first need to ensure there is an understanding on what instruments are available for adjusting and guiding the behaviour of machines.

The work group discussed the basic concepts in normatively regulating the behaviour of machines and artificial agents, as well as the basic approaches. machine:

- Functional level: the machine is either constrained to operating in an environment in which norm violation cannot happen. Of course unintended norm violations can never be ruled out entirely.
- Normative level: the machine is provided with norms that reduce the action space to those actions that comply with the norms (most likely, most of the time).
- Value level: the machine is provided with values with which it needs do align its choice of norm-guided actions, especially when norm conflicts arise.

What intervention we choose to do depends on many factors. The aim of the group is provide a joint article that can be used as an interface to the state of the art in the field of normative reasoning withing multi-agent systems.

5 Open problems

5.1 Is HOL (as a metalogic) all we need for flexible normative reasoning?

Christoph Benzmüller (Universität Bamberg, DE)

Joint work of Christoph Benzmüller, Xavier Parent, Leendert W. N. van der Torre, David Fuenmayor, Aleaxander Steen, Geoff Sutcliffe
Main reference Christoph Benzmüller, Xavier Parent, Leendert W. N. van der Torre: "Designing normative theories for ethical and legal reasoning: LogiKEy framework, methodology, and tool support", Artif. Intell., Vol. 287, p. 103348, 2020.
URL https://doi.org//10.1016/j.artint.2020.103348
Main reference Christoph Benzmüller, David Fuenmayor, Alexander Steen, Geoff Sutcliffe: "Who Finds the Short Proof?", Logic Journal of the IGPL, p. jzac082, 2023.
URL https://doi.org/10.1093/jigpal/jzac082

In previous work we have shown that classical higher-order logic (HOL), when used as a metalogic, enables (shallow) semantic embeddings of various state-of-the-art logics for normative reasoning. To this end, the logico-pluralistic LogiKEy [1] methodology and framework has been developed to support both metalogical studies of logics for normative reasoning [2] and their applications [3].

In this talk I summarise these developments and ask the obvious question: Is HOL already all we need to support flexible normative reasoning on computers? Or are there logics for normative reasoning that cannot be addressed by the LogiKEy approach?

We also briefly address typical arguments against HOL, namely that undecidability and complexity considerations militate against its use. With reference to very recent practical work on speeding up proofs in HOL [4], we will take a partially contrary position.

References

1 C. Benzmüller, X. Parent, L. van der Torre. Designing Normative Theories for Ethical and Legal Reasoning: LogiKEy Framework, Methodology, and Tool Support. Artificial Intelligence, 287: 103348. 2020. http://doi.org/10.1016/j.artint.2020.103348 (Preprint: https://www.researchgate.net/publication/342146653)
2 X. Parent, C. Benzmüller. Automated Verification of Deontic Correspondences in Isabelle/HOL – First Results. In Benzmüller, C., & Otten, J., editor(s), ARQNL 2022: Automated Reasoning in Quantified Non-Classical Logics. Proceedings of the 4th International Workshop on Automated Reasoning in Quantified Non-Classical Logics (ARQNL 2022) affiliated with the 11th International Joint Conference on Automated Reasoning (IJCAR 2022). Haifa, Israel, August 11, 2022, volume 3326, pages 92-108, 2023. CEUR Workshop Proceedings, CEUR-WS.org. https://ceur-ws.org/Vol-3326/
3 D. Fuenmayor, C. Benzmüller. Normative Reasoning with Expressive Logic Combinations. In De Giacomo, G., Catala, A., Dilkina, B., Milano, M., Barro, S., Bugarín, A., & Lang, J., editor(s), ECAI 2020 – 24th European Conference on Artificial Intelligence, June 8-12, Santiago de Compostela, Spain, volume 325, of Frontiers in Artificial Intelligence and Applications, pages 2903-2904, 2020. IOS Press. http://doi.org/10.3233/FAIA200445
4 C. Benzmüller, D. Fuenmayor, A. Steen, G. Sutcliffe. Who Finds the Short Proof? Logic Journal of the IGPL. 2023. http://doi.org/10.1093/jigpal/jzac082 (Preprint: https://www.researchgate.net/publication/367464450)

Participants

- Guillaume Aucher
University of Rennes, FR
- Kevin Baum
DFKI – Saarbrücken, DE
- Christoph Benzmüller
Universität Bamberg, DE
- Jan M. Broersen
Utrecht University, NL
- Pedro Cabalar
University of Coruña, ES
- Ilaria Canavotto
University of Maryland – College Park, US
- Agata Ciabattoni
TU Wien, AT
- Célia da Costa Pereira
Université Côte d'Azur – Sophia Antipolis, FR
- Mehdi Dastani
Utrecht University, NL
- Louise A. Dennis
University of Manchester, GB
- Frank Dignum
University of Umeå, SE
- Virginia Dignum
University of Umeå, SE
- Huimin Dong
Sun Yat-Sen University – Zhuhai, CN
- Thomas Eiter
TU Wien, AT
- Eleonora Giunchiglia
TU Wien, AT
- Guido Governatori
Tarragindi, AU
- John F. Horty
University of Maryland – College Park, US
- Joris Hulstijn
University of Luxembourg, LU
- Aleks Knoks
University of Luxembourg, LU
- Emiliano Lorini
CNRS – Toulouse, FR
- Bertram F. Malle
Brown University – Providence, US
- Réka Markovich
University of Luxembourg, LU
- Eric Pacuit
University of Maryland – College Park, US
- Xavier Parent
TU Wien, AT
- Bijan Parsia
University of Manchester, GB
- Adrian Paschke
FU Berlin, DE
- Henry Prakken
Utrecht University, NL
- Antonino Rotolo
University of Bologna, IT
- Ken Satoh
National Institute of Informatics – Tokyo, JP
- Marija Slavkovik
University of Bergen, NO
- Kai Spiekermann
London School of Economics, GB
- Christian Straßer
Ruhr-Universität Bochum, DE
- Leon van der Torre
University of Luxembourg, LU

Report from Dagstuhl Seminar 23152

Secure and Efficient Post-Quantum Cryptography in Hardware and Software

Thomas Pöppelmann*[1], Sujoy Sinha Roy*[2], and Ingrid Verbauwhede*[3]

1 **Infineon Technologies AG – Neubiberg, DE.** `thomas.poeppelmann@infineon.com`
2 **TU Graz, AT.** `sujoy.sinharoy@iaik.tugraz.at`
3 **KU Leuven, BE.** `ingrid.verbauwhede@esat.kuleuven.be`

Abstract

NIST recently announced the winners of its post-quantum cryptography (PQC) standardization process and outlined the next steps in its ongoing standardization efforts. With fewer algorithms now in the focus of the cryptographic community, the time has come to intensify the investigation of efficiency and physical security aspects of PQC algorithms. This is required to enable PQC in real-life applications and to provide feedback to NIST and submitters before final standardization. To allow widespread adoption, the implementation of PQC in current microchip technologies must be possible within application- or platform-specific constraints such as area, memory, time, power, and energy budgets. Furthermore, more and more PQC use-cases require resistance to physical attacks like power analysis.

The primary aim of this Dagstuhl Seminar was to initiate deeper investigations into secure and efficient implementations of PQC on hardware and hardware/software codesign platforms. In this direction, the seminar brought together researchers in theoretical cryptology, applied cryptography, cryptographic hardware and software systems, and physical security. During the seminar, participants identified new challenges and research directions in PQC, exchanged thoughts and ideas, and initiated collaborations on researching secured and efficient design methodologies for PQC.

Seminar April 10–13, 2023 – https://www.dagstuhl.de/23152

2012 ACM Subject Classification Security and privacy → Public key (asymmetric) techniques; Security and privacy → Hardware security implementation; Security and privacy → Hardware attacks and countermeasures

Keywords and phrases Post-quantum cryptography, secure hardware and software, cryptographic implementations, side-channel attacks, fault attacks, countermeasures against attacks

Digital Object Identifier 10.4230/DagRep.13.4.24

* Editor / Organizer

Secure and Efficient Post-Quantum Cryptography in Hardware and Software, *Dagstuhl Reports*, Vol. 13, Issue 4, pp. 24–39

Editors: Thomas Pöppelmann, Sujoy Sinha Roy, and Ingrid Verbauwhede

DAGSTUHL REPORTS Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Executive Summary

Thomas Pöppelmann
Sujoy Sinha Roy
Ingrid Verbauwhede

Our present-day public-key infrastructures primarily rely on RSA and elliptic curve cryptography (ECC). In case a powerful quantum computer is built in the near future, these public-key infrastructures will become completely insecure. Post-quantum cryptography (PQC) aims at developing new cryptographic protocols that will remain appropriately secure even after powerful quantum computers are built.

Even if powerful quantum computers are still far out, replacement algorithms for public-key algorithms need to be developed and implemented now. These algorithms must show appropriate cryptographic security, i.e., resistant to the attacks from quantum and classical computers. On top, their implementations need to be efficient in current microchip technologies, implementable within the constrained area, time, power, and energy budgets. This is very important to enable PQC-based protection of information processed by (battery-powered) Internet of things (IoT) devices or smart cards. At the same time, more and more use-cases require resistance to physical attacks. When an attacker has physical access to a device, the attacker may try to manipulate or observe it during cryptographic operations. The most common physical attacks are side-channel and fault attacks that usually aim to extract a secret key.

The existing PQC algorithms are classified into five categories depending on their underlying hard problems: lattice-based, multivariate polynomial-based, hash-based, code-based, and supersingular isogeny-based. Of them, lattice-based PQC is currently the frontrunner as evident from the fact that the majority of the PQC candidate schemes that were submitted to NIST's Post Quantum Cryptography Standardisation project are lattice-based. A significant volume of research has been performed on studying the security, performance, and application aspects of lattice-based PQC and even more narrowly focused on the algorithms submitted to the NIST call.

There is a need to have a diverse set of algorithms for post-quantum public-key cryptography. One main concern is security and risk management: if a specific class of PQC becomes weaker or is even considered broken in the advent of new cryptanalysis, then there must be other reliable classes of PQC that will offer high security. Indeed the 4th round of NIST's PQC standardization, which will start at the end of the 3rd round, will aim at broadening the set of PQC algorithms. Furthermore, in this direction, NIST indicated that a new call for proposals for PQC signature algorithms (focusing on non-lattice-based algorithms) is planned with a deadline in 2023. Besides the security aspects, each class of PQC has its own advantages. For example, code-based key agreement schemes have small ciphertexts and could be useful in applications where the public keys are known. The isogeny-based key agreement scheme SIKE has the smallest public-key and a small ciphertext size but relatively low performance. In the last few years, several new isogeny-based signature schemes have been developed with small key and signature sizes. Hash-based signature schemes have security guarantees based on hash functions and they have the advantage of (re)using a hash hardware module if the hardware platform has it. Multivariate signature schemes offer fast signing and verifying and very short signatures.

This Dagstuhl Seminar focused on answering the following questions in the context of post-quantum cryptography.

- **Efficiency and correct metrics:** Depending on the application, efficiency can be the area or memory size, throughput or latency, power and energy, or a combination of them. Can we have tailored implementations to satisfy one or several such metrics?
- **HW/SW Co-design:** The right form of interaction of a CPU with HW-based post-quantum acceleration needs to be determined: Options are instruction set extension or usage of domain-specific co-processors. How to determine the splitting of computation tasks between HW and SW?
- **Agility and reuse:** How can complex HW accelerators and controlling SW be reused? For example, can a compact HW accelerator be reused for a high throughput version? And how easy can different processing units, such as polynomial arithmetic or hash modules, support multiple schemes?
- **Physical attacks:** For many use-cases, PQC implementations need to be resistant to side-channel and fault-based attacks. Are low overhead countermeasures feasible? Shall countermeasures be implemented in HW or SW? Can we exploit the mathematical properties of some PQC algorithms to derive low-overhead countermeasures?
- **Proactive security:** Can we construct new PQC algorithms in such a way that they become more resistant to physical attacks and more efficient in HW and SW by design?

To find answers to the above-mentioned questions, the following workgroups were formed:

1. Efficient implementation aspects of PQC
2. Physical security aspects of PQC
3. Theoretical aspects of PQC
4. Application and migration

The time table of the seminar is shown in Fig. 1.

Dagstuhl Seminar 23152 Plan

Time (April)	Monday	Tuesday	Wednesday	Thursday
7:30 - 8:45		Breakfast	Breakfast	Breakfast
9:00 - 10:15		Introduction, goals, and organization. What are you looking for?	Small workgroups	Discussions on research challenges and collaboration ideas.
10:15 - 10:45		Coffee break	Coffee break	Coffee break
10:45 - 12:15		Talk (40 min) on HW/SW Acceleration of Lattice-Based Cryptography (Speaker: Tim Fritzmann) Q&A, discussions, notes.	Short report on work groups. Talk (40 min) on security metrics and certification for PQC (Speaker: Melissa Rossi). Q&A, discussions, notes.	Report of small workgroups, followed by plenary
12:15 - 14:00		Lunch break	Lunch break	Lunch break
14:00 - 15:30		Talk (40 min) on new problems in isogeny crypto. (Speaker: Christophe Petit) Q&A, discussions, notes.	Social activity (hiking and group work)	Conclusion and farewell (30 min)
15:30 - 16:00		Coffee break	Coffee break	
16:00-17:30	Arrival and Dinner	Talk (40 min) on achieving crypto agility in HW/SW. (Speaker: Matthias Kannwischer) Q&A, discussions, notes.	Discussion, ranking of most challenging topics for research.	
18:00 - 19:00		Dinner	Dinner	
20:00		Evening activity		

Figure 1 Seminar Plan.

2 Table of Contents

3 Overview of Talks

3.1 HW/SW Acceleration of Lattice-Based Cryptography

Tim Fritzmann (Infineon Technologies AG – Neubiberg, DE)

Lattice-based cryptography introduces new mathematical operations that are hard to compute on devices with a low computing power. Therefore, hardware accelerators can be used to meet performance and energy requirements. While previous works focused on standalone hardware solutions for the complete cryptographic scheme, the current trend is to use hardware/software codesigns in order to increase the flexibility of the design. In this context, two types of accelerators can be developed. Loosely coupled accelerators are suitable for large tasks with a low data transfer between main processor and accelerator. In contrast, tightly coupled accelerators are directly integrated into the main processor and avoid any complex bus communication. Experiments have shown that this type of accelerator leads to fast and flexible implementations of lattice-based cryptography.

3.2 New Problems In Isogeny-based Cryptography

Christophe Petit (Université libre de Bruxelles, BE & University of Birmingham, GB)

We give an overview of isogeny-based cryptography, including the main underlying hard problems and existing protocols. We then describe open problems in the field, with a special focus on problems relevant for hardware implementations.

3.3 Implementing the NIST PQC standards on microcontrollers

Matthew Kannwischer (Academia Sinica – Taipei, TW)

In July 2022, the US National Institute of Standards and Technology (NIST) announced the first set of post-quantum schemes to be standardized: Kyber, Dilithium, Falcon, and SPHINCS+. In this talk, I will present the state-of-the-art of those to-be-standardized schemes on the Arm Cortex-M4 which is NIST's primary microcontroller optimization target. I will present the most recent results of the benchmarking framework pqm4 for all four schemes. While for Falcon and SPHINCS+ there has been very little progress in implementation performance lately, recent improvements exist to the speed and memory consumption of Kyber and Dilithium. I will present those new implementation techniques. I will also outline new challenges for implementations on larger Arm processors and the upcoming NIST PQC signature on-ramp in particular with respect to two digital signature submissions that I am involved in: Oil-and-Vinegar and MAYO.

3.4 ANSSI recommendations and brainstorm ideas

Melissa Rossi (ANSSI – Paris, FR)

I presented the PQC transition strategy in France. The first aspect of this transition is the mandatory use of hybridation mode for PQC algorithms. I presented several modes that seem possible solutions. The second aspect is a list of potential good PQC algorithms: Kyber, FrodoKEM, Dilithium, Falcon, XMSS, LMS and SPHINCS+. Finally, I described the certification strategy in France and how it will handle post-quantum products. I concluded the talk with interesting open questions on side-channel and lattices.

4 Working Groups

4.1 Efficient implementations of PQC

Ingrid Verbauwhede (KU Leuven, BE)
Bo-Yin Yang (Academia Sinica – Taipei, TW)
Erkay Savaş (Sabanci University – Istanbul, TR)
Patrick Karl (TU München, DE)
Ahmet Can Mert (TU Graz, AT)

Our work group focused on the implementation aspects of PQC. Initially, we started with a brainstorming session to list some of the important possible research topics/directions related to efficient PQC implementations. After our discussion, we identified the following research questions.

- Implementation aspects of lattice, code, multivariate and isogeny based schemes
- Implementation optimizations targeting memory-constrained devices
- Crypto-agility in HW/SW
- Exploring synergies in different categories of PQC
- New computing paradigms for PQC (i.e., in-memory and approximate computing)
- Automatic tooling, correctness and formal verification
- Standardization of PQC-related operations into RISC-V (i.e., modular arithmetic, butterfly operation, vector operations etc.)

After the initial discussions, the following two main directions were brainstormed by this work group.

4.1.1 Formal verification of PQC implementations

Two approaches for automating the correctness verification process (formal methods) are discussed, model checking by SAT solver and proof assistant. The model checking defines what should or should not happen and then evaluates the given program. The result is either satisfiable, unsatisfiable or non-determine. Domain specific languages can also help in the verification process significantly. We discussed the formal verification tools like CryptoLine [1] which is a language for the verification of low-level mathematical constructions.

Demo. Dr. Bo-Yin Yang provided a demo on how to use CryptoLine and explained basic working principles. CryptoLine uses low-level instruction models for specific micro-controllers provided by Jasmin. It translates each low-level instruction into one or several CryptoLine instructions. Even for complex programs, it takes only couple of minutes for CryptoLine to perform correctness verification.

CryptoLine shows a significant improvement towards tooling/automation of correctness verification. However, there are still challenges such as handling of noise sampling and decryption failures. Verification of floating-point based implementations is also very challenging due to operations like rounding, truncation and overflow. Another important challenge is translating these verification approaches to hardware implementations of PQC.

4.1.2 Exploring synergies in different categories of PQC

NIST finalized its standardization process and announced four candidates (one PKE and three DS) to be standardized. Besides, different cybersecurity agencies have suggested a gradual transition to PQC and some of them selected deployment of a different scheme than NIST's standardization process winners. It is also possible that some of the non-standardized schemes still can find use in some specific platforms and applications. This shows that unified implementations supporting several schemes will be required soon. To that end, it has important significance to find and explore synergies in different PQC schemes.

Our work group started with identifying common arithmetic operations in different PQC categories such as lattice-based and code-based cryptography. Our initial investigation showed that hashing, modular arithmetic and number theoretic transform (NTT) are common operations in most PQC schemes. Our discussion further led to the following research questions/directions.

- Can approximate computing (i.e., allowing errors in computation at the expense of increased failure rate) lead to *super low power* applications?
- Use of erroneous multipliers to improve power consumption.
- Design of hardware modules that lead to some increase in failure rate but allow to reuse them for other schemes, e.g. common multiplier for SABER/Kyber.
- Modelling/Verifying the aforementioned approaches.
- Exploring schemes (mostly non lattice-based) in NIST's additional digital signature call. Since this is very new, this is very unexplored in terms of HW/SW implementations.

Conclusions from this workgroup. (*i*) Recent efforts in formal verification of PQC are promiisng and important; however, there are challenges/limitations such as floating-point arithmetic. (*ii*) This work group will continue collaborative study for exploring synergies in different PQC schemes.

Open problems from this workgroup. (*i*) Translating existing correctness verification approaches/tools to hardware implementations of PQC.

References

1 Yu-Fang Chen, Chang-Hong Hsu, Hsin-Hung Lin, Peter Schwabe, Ming-Hsien Tsai, Bow-Yaw Wang, Bo-Yin Yang, Shang-Yi Yang. *Verifying Curve25519 Software.* ACM SIGSAC Conference on Computer and Communications Security (CCS), pp. 299-309, ACM, 2014.

4.2 Physical security aspects of PQC

Aydin Aysu (North Carolina State University – Raleigh, US)
Ayesha Khalid (Queen's University of Belfast, GB)
Gaetan Cassiers (TU Graz, AT)
Melissa Rossi (ANSSI – Paris, FR)
Peter Pessl (Infineon Technologies AG – Neubiberg, DE)
Prasanna Ravi (Temasek Labs, Singapore & Nanyang Technological University, Singapore, SG)
Simona Samardjiska (Radboud University Nijmegen, NL)
Thomas Eisenbarth (Universität Lübeck, DE)

Our workgroup mainly foccused on the broad topic of physical attacks such as SCA and FIA of PQC schemes. We started by laying out some of the open-problems, that need to be addressed to improve our understanding of the threat of SCA and FIA, as well as implementing efficient PQC designs in a manner resistant to physical attacks. After intense deliberation and discussion, we were able to identify several research questions and open problems that can be split into three broad categories.

- Evaluation of side-channel leakage for practical implementations of PQC schemes.
- Efficient Countermeasures against SCA for Post-Quantum Cryptographic schemes
- Identification of new SCA and FIA on PQC schemes

In the following, we briefly explain the several research questions and discussions that emerged out of the discussion in our work group.

4.2.1 Evaluation of Side-Channel Leakage for Post-Quantum Cryptographic Schemes

We will soon be witnessing wide-scale adoption of lattice-based schemes on embedded devices, and these devices have to be evaluated based on different security certification standards such as FIPS 140-3 [2], Common Criteria [3]. In order to perform side-channel leakage evaluation of PQC schemes, we are not aware of the concrete set of tests that need to be done to certify a given PQC hardware or software against side-channel analysis. Thus, it can be an interesting research direction from the point of view of security certification of PQC HW and SW implementations.

Some of the specific open-problems that we considered during our discussions are as follows:

1. What are the exhaustive set of tests required to test side-channel leakage from all the operations within the decapsulation procedure of IND-CCA secure Key Encapsulation Mechanisms (KEMs). The first challenge towards devising an exhaustive set is the fact that the decapsulation procedure contains three different operations (i.e.) decryption, re-encryption and ciphertext comparison. Thus, it is necessary to device separate tests to test for leakage from each of these three operations separately. While it does not appear to be extremely difficult to arrive at such exhaustive tests for individual post-quantum KEMs, it needs to be analyzed whether they make up the exhaustive set of tests that are sufficient to prove existence of leakage or otherwise.

2. While leakage evaluation tests can help us detect or test for leakage of sensitive variables, it does not necessarily indicate the possibility to perform key recovery. Thus, it is also interesting to ponder upon development of novel techniques that can map available leakage to the most efficient key recovery attack. We are probably not looking at automatic discovery of new attacks based on existing vulnerabilities, but probably estimate the effort required to mount known attacks provided the calculated leakage. This could be particularly interesting for the case of single trace attacks [4, 5], where we can ask the question what is the minimum amount of leakage required to mount a given single trace attack, without explicitly performing the attack itself.
3. Can we develop techniques that can ascertain if certain types of PQC schemes are more difficult to attack through SCA/FIA compared to others? This is for instance, a pertinent question that came up during the NIST PQC process, where NIST was particularly interested in factors that differentiated the different lattice-based schemes in the context of side-channel analysis, given that there were several lattice-based schemes in the NIST PQC standardization process. Development of such analysis techniques can potentially enable us to build leakage resilient schemes that are inherently resistant to SCA [6, 8, 7].

4.2.2 Efficient Countermeasures against Side-Channel Attacks for PQC schemes

1. Masking countermeasures for PQC schemes typically involve a large amount of randomness, especially because they involve computation over large polynomials, matrices or vectors spanning dimensions of the order of a few hundred to few thousand. Obtaining access to a large amount of high quality randomness typically required for masking schemes is particularly challenging, especially when considering constrained embedded devices, which are especially required to be protected against side-channel attacks. These challenges for masking countermeasures gives rise to the following questions.
   a. What is the effect of reusing randomness in masking schemes, tailor made for specific PQC schemes to reduce the true performance overhead of masking countermeasures? This will reduce the randomness requirement, thereby reducing the true performance overhead of masking countermeasures.
   b. What is the impact of using randomness of bad quality in masked implementations of PQC schemes?
2. Masking countermeasures are usually considered to be very expensive, as they have shown to incur a significant overhead in runtime in several prior works [9, 10, 11]. We also observe that PQC schemes involve computation over large polynomials, matrices or vectors spanning dimensions of the order of a few hundred to few thousand. Thus, it is interesting to contemplate use of shuffling countermeasures, exploiting the large dimensions of elements used in PQC based schemes. This is particularly relevant in scenarios where that an attacker is in a restricted setting, limiting him/her with access to "N" traces for a given secret key. In such scenarios, whether shuffling alone is sufficient to provide provable security?
3. Code-based schemes are currently in the spotlight as we expect a code-based scheme to be standardized in the fourth round. Interestingly, there are no systematic measures for SCA protection of these schemes. The decoding algorithms are notoriously difficult to even do in constant time, and the question is whether we can use masking techniques in the decoding algorithms – for instance Reed Muller, Reed Solomon Decoder in HQC, Berlekamp-Massey in Classic McEliece.

4.2.3 Identification of new SCA and FIA on PQC schemes

1. Existing side-channel attacks on lattice-based cryptographic schemes either require the knowledge of inputs/outputs, or require to control the inputs to the DUT [12]. However, it is possible in certain scenarios that the attacker does not directly obtain access to the I/O of the DUT. In such a setting, the attacker only has access to the side-channel traces, and it begets the question if an attacker can still perform key recovery without knowledge of the DUT's inputs/outputs. It is also interesting to explore for which post-quantum schemes this scenario leads to meaningful, exploitable leakage.
2. Several prior works have shown that an attacker can craft malicious inputs to the decapsulation procedure of KEMs to amplify the side-channel leakage of the secret key for efficient key recovery attacks. One of the main downsides of these attacks is that the malformed ciphertexts used for the attack, can be detected with a very high probability. Such chosen-ciphertext attacks using malformed ciphertexts have also been used to target other PQC schemes as well [1]. This makes it natural for a designer to implement a simple protection: refresh the secret key every time he/she observes a decapsulation failure, since decapsulation failures for valid ciphertexts occur with negligible probability. This raises a natural question on whether "it is possible to perform chosen-ciphertext attacks on lattice-based schemes with valid ciphertexts?". This represents a more stealthy approach towards chosen-ciphertext attacks, as valid ciphertexts cannot be detected as malicious by the decapsulation procedure, as they do not trigger decapsulation failures.

Conclusions from this workgroup. During the course of our discussions, we identified several open problems along three axes – SCA/FIA based attacks, Efficient SCA/FIA Countermeasures and the need for comprehensive SCA Leakage Evaluation Techniques of PQC schemes. However, we identified two key components that could foster further research on SCA/FIA of PQC schemes.

1. Development of an open-source implementation framework that allows to run SCA experiments on PQC schemes implemented on commonly used microcontrollers. In this respect, we discussed about the possibility of extrapolating the pqm4 library [13] and integrate with the open-source Chipwhisperer platform [14]. The framework should be developed in such a way that it serves as a library of SCA protected implementations of PQC schemes, and allows for testing for leakage in critical components within implementations of PQC schemes.
2. Development of an open-source database of side-channel traces of implementations of PQC schemes, which can motivate the community towards developing attack techniques as well as novel leakage detection techniques targeting the open-source trace database. While similar open-source side-channel trace database are available for symmetric ciphers such as the ASCAD database for AES [15], we are not aware of similar works for PQC schemes.

Open problems from this workgroup. (i) Development of novel leakage evaluation techniques for PQC schemes, that enables to certify a given PQC implementation as secure or not. (ii) Investigation of the quality of randomness used in masked implementations of PQC schemes. (iii) Development of efficient masking schemes for code-based schemes (iv) Development of novel SCA/FIA for new attack scenarios such as a) Use of Valid Ciphertexts and b) Blind SCA/FIA that work without knowledge of inputs/outputs. (v) Development of open-source implementation framework that enables SCA/FIA evaluation of PQC implementations.

References

1 Ravi, Prasanna, and Sujoy Sinha Roy. *Side-channel analysis of lattice-based PQC candidates.* In Round 3 Seminars, NIST Post Quantum Cryptography. 2021.

2 Schaffer, Kim. *FIPS 140-3 Derived Test Requirements (DTR): CMVP Validation Authority Updates to ISO/IEC 24759. No. NIST Special Publication (SP) 800-140 (Draft).* National Institute of Standards and Technology, 2019.

3 Boswell, Tony. *Security evaluation and common criteria.* In Secure Smart Embedded Devices, Platforms and Applications, pp. 407-427. New York, NY: Springer New York, 2013.

4 Primas, Robert, Peter Pessl, and Stefan Mangard. *Single-trace side-channel attacks on masked lattice-based encryption.* In Cryptographic Hardware and Embedded Systems-CHES 2017: 19th International Conference, Taipei, Taiwan, September 25-28, 2017, Proceedings, pp. 513-533. Springer International Publishing, 2017.

5 Pessl, Peter, and Robert Primas. *More practical single-trace attacks on the number theoretic transform.* In Progress in Cryptology–LATINCRYPT 2019: 6th International Conference on Cryptology and Information Security in Latin America, Santiago de Chile, Chile, October 2–4, 2019, Proceedings 6, pp. 130-149. Springer International Publishing, 2019.

6 Clément Hoffmann, Benoît Libert, Charles Momin, Thomas Peters, and François-Xavier Standaert. *POLKA: Towards Leakage-Resistant Post-quantum CCA-Secure Public Key Encryption.* In Public-Key Cryptography – PKC 2023

7 Jan-Pieter D'Anvers, Emmanuela Orsini, and Frederik Vercauteren. *Error Term Checking: Towards Chosen Ciphertext Security without Re-encryption.* In Proceedings of the 8th ACM on ASIA Public-Key Cryptography Workshop – APKC '21

8 Melissa Azouaoui, Yulia Kuzovkova, Tobias Schneider, Christine van Vredendaal, *Post-Quantum Authenticated Encryption against Chosen-Ciphertext Side-Channel Attacks.* In IACR Transactions on Cryptographic Hardware and Embedded Systems (TCHES) 2023

9 Bos, Joppe W., Marc Gourjon, Joost Renes, Tobias Schneider, and Christine Van Vredendaal. *Masking kyber: First-and higher-order implementations.* IACR Transactions on Cryptographic Hardware and Embedded Systems (2021): 173-214.

10 Heinz, Daniel, Matthias J. Kannwischer, Georg Land, Thomas Pöppelmann, Peter Schwabe, and Daan Sprenkels. *First-order masked Kyber on ARM Cortex-M4.* Cryptology ePrint Archive (2022).

11 Bronchain, Olivier, and Gaëtan Cassiers. *Bitslicing arithmetic/boolean masking conversions for fun and profit: with application to lattice-based kems.* IACR Transactions on Cryptographic Hardware and Embedded Systems (2022): 553-588.

12 Ravi, Prasanna, Anupam Chattopadhyay, Jan Pieter D'Anvers, and Anubhab Baksi. *Side-channel and fault-injection attacks over lattice-based post-quantum schemes (Kyber, Dilithium): Survey and new results.* ACM Transactions on Embedded Computing Systems (2022).

13 Kannwischer, Matthias J., Joost Rijneveld, Peter Schwabe, and Ko Stoffelen. *pqm4: Testing and Benchmarking NIST PQC on ARM Cortex-M4.* (2019).

14 O'flynn, Colin, and Zhizhang Chen. *Chipwhisperer: An open-source platform for hardware embedded security research.* In Constructive Side-Channel Analysis and Secure Design: 5th International Workshop, COSADE 2014, Paris, France, April 13-15, 2014. Revised Selected Papers 5, pp. 243-260. Springer International Publishing, 2014.

15 Benadjila, Ryad, Emmanuel Prouff, Rémi Strullu, Eleonora Cagli, and Cécile Dumas. *Deep learning for side-channel analysis and introduction to ASCAD database.* Journal of Cryptographic Engineering 10, no. 2 (2020): 163-188.

4.3 Theoretical aspects of PQC

Jan-Pieter D'Anvers (KU Leuven, BE)
Andrea Basso (University of Bristol, GB)
Thomas Pöppelmann (Infineon Technologies – Neubiberg, DE)
Thomas Prest (PQShield – Paris, FR)
Tobias Schneider (NXP Semiconductors – Gratkorn, AT)
Christophe Petit (UL – Brussels, BE & University of Birmingham, GB)
Sujoy Sinha Roy (TU Graz, AT)

Our workgroup focused on the cost of protecting implementations against side-channel attacks. More specifically, we looked at what design choices can make cryptographic schemes easier to protect against these side-channel attacks. Our brainstorming resulted in three main topics: replacing the FO transformation, improvements to the NIST standard Kyber, and new methods in masking.

The FO transformation is a widely used method to secure schemes actively. However, it comes at a great cost, making ciphertext decryption 2 to 3 times more expensive and making masking harder due to its nonlinearity. We looked at three methods to replace the FO transformation for LWE-based schemes: POLKA [1] and ETC [2] and an ID-based proposal [3]. All these methods are interesting but come with specific preconditions. We looked at the possibility of a more general method and the limitations of what is possible in this space.

Regarding improvements to Kyber, we discussed the possibility of arithmetic hash functions, as other Kyber operations are also arithmetic which could reduce the need for costly conversions between arithmetic and Boolean domains. However, we concluded that this is not trivial due to the inherent conversion between the arithmetic and Boolean domain during the decoding of the message.

In the masking domain, we discussed different security models for probing, which would better mimic existing attacks and subsequently could result in more efficient masking. We also discussed the importance of reducing the randomness cost of masking algorithms. We concluded that this should be a more prominent factor in designing and evaluating these countermeasures. As for existing countermeasures, we looked at improvement possibilities. Notably, the Kavach [4] implementation could benefit from a more efficient Boolean masking of the carry, and the one-hot conversion [5] could benefit from the fact that intermediate variables have constant hamming weight and thus might be better protected against hamming weight leakage.

Conclusions from this workgroup. Using the existing mechanisms, the cost of implementing side-channel countermeasures is quite high, especially when higher order masking are considered. More research is needed.

Open problems from this workgroup. The workgroup will continue working on a concrete idea to replace the FO transformation with a different transformation. The expected result is a shift in cost from decryption to encryption. Other open problems are the improvements of the Kavach, and the one-hot conversion are not planned but left as interesting future work.

References

1 Clément Hoffmann, Benoît Libert, Charles Momin, Thomas Peters, and François-Xavier Standaert. *POLKA: Towards Leakage-Resistant Post-quantum CCA-Secure Public Key Encryption*. In Public-Key Cryptography – PKC 2023

2 Jan-Pieter D'Anvers, Emmanuela Orsini, and Frederik Vercauteren. *Error Term Checking: Towards Chosen Ciphertext Security without Re-encryption*. In Proceedings of the 8th ACM on ASIA Public-Key Cryptography Workshop – APKC '21

3 Melissa Azouaoui, Yulia Kuzovkova, Tobias Schneider, Christine van Vredendaal, *Post-Quantum Authenticated Encryption against Chosen-Ciphertext Side-Channel Attacks*. In IACR Transactions on Cryptographic Hardware and Embedded Systems (TCHES) 2023

4 Aikata Aikata and Andrea Basso and Gaetan Cassiers and Ahmet Can Mert and Sujoy Sinha Roy, *Kavach: Lightweight masking techniques for polynomial arithmetic in lattice-based cryptography*. In IACR Transactions on Cryptographic Hardware and Embedded Systems (TCHES) 2023

5 Jan-Pieter D'Anvers *One-Hot Conversion: Towards Faster Table-Based A2B Conversion*. In Advances in Cryptology – EUROCRYPT 2023

4.4 Application and migration

Thomas Eisenbarth (Universität Lübeck, Gernany)
Melissa Rossi (ANSSI – Paris, FR)
Rainer Steinwandt (University of Alabama in Huntsville, US)
Marc Stöttinger (Hochschule RheinMain – Wiesbaden, DE)

After initial brainstorming, the discussion focused on the transition phase from traditional public-key solutions to post-quantum solutions. While there is no consensus on the use of hybrid systems as an intermediate step, the interest in such solutions is significant. A common building block for hybrid designs is the use of some form of combiner that is realized with the help of a key derivation function. Depending on the application context, contributing key material may involve a pre-shared key, the outcome of a traditional key establishment, a shared secret from a post-quantum solution, key material derived from a quantum key distribution protocol. Including additional material in the key derivation may be desirable.

We looked at different approaches considered for combining key material from different sources, e.g., by ANSSI (France) [1] and by Germany's Federal Office for Information Security [2]. Aligning with the focus of this Dagstuhl Seminar, our main interest was on possible side-channel vulnerabilities when combining keys from different sources, and we started to discuss if it is realistic and possible to mount such attacks. In addition, we had an initial discussion on attacker models and possible side-channel leakage models. We also formulated the following initial research questions on SCA in combinational functions in hybrid methods:

- What is a typical recommended combiner function and how closely are these related to PRFs, MACs, or hash functions?
- What is a suitable PRF or DualPRF regarding side-channel resistance?
- Is there literature on side-channel attacks on combiner functions used in a hybrid PQC scheme?
- Is there literature on side-channel attacks on PRFs that could serve as a starting point for studying side-channel vulnerabilities in combiner functions?

The following action items have been set up to continue research on this topic even after the Dagstuhl Seminar is over.

- Conducting a literature review will be needed to decide on the next steps.
- Set up follow-up meetings to proceed and exchange on advance with respect to SCA on combiner functions.

Conclusions from this workgroup. The topic SCA on combiner functions will be continued in a collaborative workling group after the Dagstuhl Seminar.

Open problems from this workgroup. Side-Channel Analysis on Combiner Function of Hybrid Schemes.

References

1. Mélissa Rossi. PQC Transition in France; ANSSI views. Presentation at Real World Post-Quantum Crypto, March 2023. Available at `https://www.melissarossi.fr/wp-content/uploads/2023/04/ANSSIs-recommendations-on-the-migration-plan.pdf`.
2. Federal Office for Information Security. Quantum-safe cryptography – fundamentals, current developments and recommendations, May 2022. Available at `https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/Brochure/quantum-safe-cryptography.html`.

Participants

- Aydin Aysu
North Carolina State University – Raleigh, US
- Andrea Basso
University of Bristol, GB
- Gaetan Cassiers
TU Graz, AT
- Jan-Pieter D'Anvers
KU Leuven, BE
- Thomas Eisenbarth
Universität Lübeck, DE
- Tim Fritzmann
Infineon Technologies AG – Neubiberg, DE
- Mike Hamburg
Rambus – Vught, NL
- Matthias Kannwischer
Academia Sinica – Taipei, TW
- Patrick Karl
TU München, DE
- Ayesha Khalid
Queen's University of Belfast, GB
- Ahmet Can Mert
TU Graz, AT
- Peter Pessl
Infineon Technologies AG – Neubiberg, DE
- Christophe Petit
UL – Brussels, BE & University of Birmingham, GB
- Thomas Pöppelmann
Infineon Technologies AG – Neubiberg, DE
- Thomas Prest
PQShield – Paris, FR
- Prasanna Ravi
Nanyang TU – Singapore, SG & Temasek Labs – Singapore, SG
- Mélissa Rossi
ANSSI – Paris, FR
- Simona Samardjiska
Radboud University Nijmegen, NL
- Erkay Savas
Sabanci University – Istanbul, TR
- Tobias Schneider
NXP Semiconductors – Gratkorn, AT
- Sujoy Sinha Roy
TU Graz, AT
- Rainer Steinwandt
University of Alabama in Huntsville, US
- Marc Stöttinger
Hochschule RheinMain – Wiesbaden, DE
- Ingrid Verbauwhede
KU Leuven, BE
- Bo-Yin Yang
Academia Sinica – Taipei, TW

Report from Dagstuhl Seminar 23161

Pushing the Limits of Computational Combinatorial Constructions

Lucia Moura*[1], Anamari Nakic*[2], Patric Östergård*[3], Alfred Wassermann*[4], and Charlene Weiß†[5]

1 **University of Ottawa, CA.** lmoura@uottawa.ca
2 **University of Zagreb, HR.** anamari.nakic@fer.hr
3 **Aalto University, FI.** patric.ostergard@aalto.fi
4 **Universität Bayreuth, DE.** alfred.wassermann@uni-bayreuth.de
5 **Universität Paderborn, DE.** chweiss@math.uni-paderborn.de

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 23161 "Pushing the Limits of Computational Combinatorial Constructions".

In this Dagstuhl Seminar, we focused on computational methods for challenging problems in combinatorial construction. This includes algorithms for construction of combinatorial objects with prescribed symmetry, for isomorph-free exhaustive generation, and for combinatorial search. Examples of specific algorithmic techniques are tactical decomposition, the Kramer-Mesner method, algebraic methods, graph isomorphism software, isomorph-free generation, clique-finding methods, heuristic search, SAT solvers, and combinatorial optimization. There was an emphasis on problems involving graphs, designs and codes, also including topics in related fields such as finite geometry, graph decomposition, Hadamard matrices, Latin squares, and q-analogs of designs and codes.

Seminar April 16–21, 2023 – https://www.dagstuhl.de/23161
2012 ACM Subject Classification Mathematics of computing → Discrete mathematics; Mathematics of computing → Mathematical software
Keywords and phrases automorphism groups, combinatorial algorithms, finite geometries, subspace designs
Digital Object Identifier 10.4230/DagRep.13.4.40

1 Executive Summary

Lucia Moura (University of Ottawa, CA)
Anamari Nakic (University of Zagreb, HR)
Patric Östergård (Aalto University, FI)
Alfred Wassermann (Universität Bayreuth, DE)

Objectives of the seminar

In this Dagstuhl Seminar, the focus was on computational methods for challenging problems in combinatorial construction. This seminar brought together researchers with different expertise: those with a deep understanding of combinatorial objects and experts in algorithm design and high-performance computing. Participants identified important problems for

* Editor / Organizer
† Editorial Assistant / Collector

codes, graphs, and designs; and discussed state-of-the-art methods for challenging problems in constructing combinatorial objects. Additionally, the seminar brought together experts on combinatorial algorithms and representatives from different scientific communities developing practical techniques for attacking general hard problems, for example, in the framework of SAT solving, integer linear programming, and optimization.

General overview of the research topic

In discrete mathematics, construction and classification of structures are core problems. Computational methods have been essential in settling important mathematical questions, such as the proof of the four color theorem (Apel and Haken, 1976) and the nonexistence of projective planes of order 10 (Lam et al., 1989). Isomorph-free exhaustive searches are quite challenging with the number of nonequivalent objects often growing exponentially with the input size. For example, the classification of Steiner triple systems of order 19 (Kaski and Östergård, 2004) involved producing a list of more than 11 billion such objects. Vast exhaustive searches are also used to establish negative results, as in the non-existence of 16-clue Sudoku puzzles (McGuire et al., 2012), celebrated by the media due to its connection with the famous puzzle. Many of the central problems, even subproblems, are (NP-)hard, but with improved algorithms and general approaches, it is still possible to handle instances with not too small parameters.

Structure of the seminar

This seminar was conducted in three main forms: plenary sessions, tutorials, and working groups. Each morning started with a plenary session, where we asked four speakers to give 60 minute lectures:

- Curtis Bright (University of Windsor, CA): *SAT + Isomorph-free Generation ...and the Quest for the Minimum Kochen–Specker System*
- Daniel Heinlein (Aalto University, FI): *Enumerating Steiner Triple Systems: Counting STS(21)s*
- Vedran Krčadinac (University of Zagreb, HR): *On higher-dimensional designs*
- Pascal Schweitzer (TU Darmstadt, DE): *Automorphisms, Isomorphisms, and Canonization: recent developments*

In the first two days of the seminar six leading experts gave 30 minute tutorials and provided input on the current limits of the area:

- Brendan McKay (Australian National University – Acton, AU): *SURGE : A fact open-source chemical graph generator*
- Gordon Royle (The University of Western Australia – Crawley, AU): *Three stories about computational combinatorics*
- Manfred Scheucher (TU Berlin, DE): *Using SAT Solvers in Combinatorics and Geometry*
- Brett Stevens (Carleton University – Ottawa, CA): *Thoughts on Computational Design Theory*
- Leo Storme (Ghent University, BE): *Computational methods in finite geometry*
- Ian M. Wanless (Monash University – Clayton, AU): *Open problems on orthogonal(ish) arrays*

In the remaining time participants were partitioned into five working groups, brainstorming and exchanging ideas:

- *Improving reliability and usability of computational projects*
- *SAT working group*
- *Isomorphism Solvers*
- *Design theory working group*
- *Tactical decompositions working group*

2 Table of Contents

3 Overview of Talks

3.1 SAT + Isomorph-free Generation ...and the Quest for the Minimum Kochen–Specker System

Curtis Bright (University of Windsor, CA)

Joint work of Zhengyu Li, Curtis Bright, Vijay Ganesh
Main reference Zhengyu Li, Curtis Bright, Vijay Ganesh: "An SC-Square Approach to the Minimum Kochen–Specker Problem, SC-Square Workshop, 2022.
URL https://cs.uwaterloo.ca/ cbright/reports/sc2-ks-preprint.pdf

I will describe a new approach for exhaustively generating combinatorial objects by combining a satisfiability (SAT) solver with an isomorph-free exhaustive generation method such as orderly generation. The SAT solver is able to limit the search to objects that satisfy given criteria, while the isomorph-free generation method ensures that the objects are generated up to isomorphism. The combined search procedure performs orders-of-magnitude faster than a pure SAT or pure computer algebraic approach, as the SAT solver tailors the search to the object in question while the isomorph-free generation avoids duplication of work when the search space is highly symmetrical.

As a motivating example, I will discuss how this approach can be applied to search for Kochen–Specker (KS) systems, an important combinatorial object arising in quantum physics. An exhaustive computer search in 2016 was able to disprove the existence of a KS system of 21 or fewer vectors. Our SAT and orderly generation approach is over 32,000 times faster than the previously used approach and has also ruled out the existence of a KS system with 22 vectors.

References

1 J. Conway, S. Kochen. *The Free Will Theorem*. Foundations of Physics, 2006.
2 S. Uijlen, B. Westerbaan. *A Kochen-Specker System Has at Least 22 Vectors*. New Generation Computing, 2016.
3 C. Bright, J. Gerhard, I. Kotsireas, V. Ganesh. *Effective Problem Solving Using SAT Solvers*. Maple Conference 2019.
4 C. Bright, K. Cheung, B. Stevens, D. Roy, I. Kotsireas, V. Ganesh. *A nonexistence certificate for projective planes of order ten with weight 15 codewords*. Applicable Algebra in Engineering, Communication and Computing, 2020.
5 C. Bright, K. Cheung, B. Stevens, I. Kotsireas, V. Ganesh. *A SAT-based Resolution of Lam's Problem*. AAAI 2021.
6 T. Junttila, M. Karppa, P. Kaski, J. Kohonen. *An adaptive prefix-assignment technique for symmetry reduction*. Journal of Symbolic Computation, 2020.
7 J. Savela, E. Oikarinen, M. Järvisalo. *Finding periodic apartments via Boolean satisfiability and orderly generation*. LPAR 2020.
8 M. Kirchweger, M. Scheucher, S Szeider. *A SAT Attack on Rota's Basis Conjecture*. SAT 2022.
9 Z. Li, C. Bright, V. Ganesh. *An SC-Square Approach to the Minimum Kochen–Specker Problem*, SC-Square Workshop, 2022.
10 C. Bright, I. Kotsireas, V. Ganesh. *When Satisfiability Checking Meets Symbolic Computation*. Communications of the ACM, 2022.

3.2 Enumerating Steiner Triple Systems: Counting STS(21)s

Daniel Heinlein (Aalto University, FI)

Joint work of Östergård, Patric R. J.
Main reference Daniel Heinlein, Patric R. J. Östergård: "Enumerating Steiner Triple Systems", CoRR, Vol. abs/2303.01207, 2023
URL https://arxiv.org/abs/2303.01207

Steiner triple systems (STSs) have been classified up to order 19. Earlier estimations of the number of isomorphism classes of STSs of order 21, the smallest open case, are discouraging as for classification, so it is natural to focus on the easier problem of merely counting the isomorphism classes. Computational approaches for counting STSs are here considered and lead to an algorithm that is used to obtain the number of isomorphism classes for order 21: 14,796,207,517,873,771.

3.3 On higher-dimensional designs

Vedran Krčadinac (University of Zagreb, HR)

Joint work of Vedran Krčadinac, Mario Osvin Pavčević, Kristijan Tabak
Main reference Vedran Krčadinac, Mario Osvin Pavčević, Kristijan Tabak: "Cubes of symmetric designs", CoRR, Vol. abs/2304.05446, 2023.
URL https://doi.org/10.48550/arXiv.2304.05446

Higher-dimensional Hadamard matrices were introduced in the 1970s by Paul Shlichta [3, 4]. In 1990, Warwick de Launey [1] developed a general framework for higher-dimensional designs of various types, including symmetric designs, Hadamard matrices, and their generalizations. In this talk I will focus on n-dimensional Hadamard matrices and symmetric designs. I will give an overview of the known constructions and present a new construction giving examples that may have inequivalent slices. The new construction was recently discovered in a joint work with Mario Osvin Pavčević and Kristijan Tabak [2]. There are many open questions in this area, including the existence of symmetric designs for some small parameters (v, k, λ) and dimensions $n \geq 3$ when they are known to exist for $n = 2$. Some of these problems could be good candidates for clever computer constructions.

References

1 W. de Launey, *On the construction of n-dimensional designs from 2-dimensional designs*, Combinatorial mathematics and combinatorial computing, Vol. 1 (Brisbane, 1989). Australas. J. Combin. **1** (1990), 67–81.
2 V. Krčadinac, M. O. Pavčević, K. Tabak, *Cubes of symmetric designs*, preprint, 2023. `http://arxiv.org/abs/2304.05446`
3 P. J. Shlichta, *Three- and four-dimensional Hadamard matrices*, Bull. Amer. Phys. Soc. **16 (8)** (1971), 825–826.
4 P. J. Shlichta, *Higher dimensional Hadamard matrices*, IEEE Trans. Inform. Theory **25** (1979), no. 5, 566–572.

3.4 SURGE : A fact open-source chemical graph generator

Brendan McKay (Australian National University – Acton, AU)

Joint work of Brendan D. McKay, Christoph Steinbeck, Mehmed Aziz Yirik
Main reference Brendan D. McKay, Mehmet Aziz Yirik, Christoph Steinbeck: "Surge: a fast open-source chemical graph generator", J. Cheminformatics, Vol. 14(1), p. 24, 2022.
URL https://doi.org//10.1186/s13321-022-00604-9

Chemical structure generators are used in cheminformatics to produce or enumerate virtual molecules based on a set of boundary conditions. The result can then be tested for properties of interest, such as adherence to measured data or for their suitability as drugs. The starting point can be a potentially fuzzy set of fragments or a molecular formula. In the latter case, the generator produces the set of constitutional isomers of the given input formula. Here we present the novel constitutional isomer generator surge based on the canonical generation path method. Surge uses the nauty package to compute automorphism groups of graphs. We outline the working principles of surge and present benchmarking results which show that surge is currently the fastest structure generator. Surge is available under a liberal open-source license.

3.5 Three stories about computational combinatorics

Gordon Royle (The University of Western Australia – Crawley, AU)

I discuss three projects in computational combinatorics that highlight issues relating to correctness, reliability and re-usability of the results obtained by such projects.

The three projects discussed are the proof of the four-colour theorem, the proof of the non-existence of a projective plane of order 10 and the construction of the catalogue of 8-element matroids.

3.6 Using SAT Solvers in Combinatorics and Geometry

Manfred Scheucher (TU Berlin, DE)

In this talk, we discuss how modern SAT solvers can be used to tackle mathematical problems. We discuss various problems to give the audience a better understanding, which might be tackled in this fashion, and which might not. Besides the naive SAT formulation further ideas are sometimes required to tackle problems. Additional constraints such as statements which hold "without loss of generality" might need to be added so that solvers terminate in reasonable time.

3.7 Automorphisms, Isomorphisms and Canonization: recent developments

Pascal Schweitzer (TU Darmstadt, DE)

Joint work of Markus Anders, Jendrik Brachter, Martin Grohe, Pascal Schweitzer, Julian Stieß, Daniel Wiebking

The Graph Isomorphism Problem, the task of computing automorphisms groups, and Canonization are closely related tasks revolving around symmetries. Indeed, the task of computing symmetries is known to be equivalent to the isomorphism problem, both theoretically and practically. In my talk I will survey two recent advances in the area of algorithmic symmetry detection and exploitation.

I will describe recent improvements in practical graph isomorphism solvers. I will also hint at new insights regarding the structure of automorphism groups of graphs subject to various restrictions. Finally, I will relate canonization algorithms of general combinatorial objects to canonization of graphs and describe new algorithmic ideas for this problem.

The talk reports on various papers including joint work with Markus Anders, Jendrik Brachter, Martin Grohe, Julian Stieß, and Daniel Wiebking.

References

1 Markus Anders, Jendrik Brachter, and Pascal Schweitzer. A characterization of individualization-refinement trees. In Hee-Kap Ahn and Kunihiko Sadakane, editors, *32nd International Symposium on Algorithms and Computation, ISAAC 2021, December 6-8, 2021, Fukuoka, Japan*, volume 212 of *LIPIcs*, pages 24:1–24:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021.

2 Markus Anders and Pascal Schweitzer. Engineering a fast probabilistic isomorphism test. In Martin Farach-Colton and Sabine Storandt, editors, *Proceedings of the Symposium on Algorithm Engineering and Experiments, ALENEX 2021, Virtual Conference, January 10-11, 2021*, pages 73–84. SIAM, 2021.

3 Markus Anders and Pascal Schweitzer. Parallel computation of combinatorial symmetries. In Petra Mutzel, Rasmus Pagh, and Grzegorz Herman, editors, *29th Annual European Symposium on Algorithms, ESA 2021, September 6-8, 2021, Lisbon, Portugal (Virtual Conference)*, volume 204 of *LIPIcs*, pages 6:1–6:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021.

4 Markus Anders and Pascal Schweitzer. Search problems in trees with symmetries: Near optimal traversal strategies for individualization-refinement algorithms. In Nikhil Bansal, Emanuela Merelli, and James Worrell, editors, *48th International Colloquium on Automata, Languages, and Programming, ICALP 2021, July 12-16, 2021, Glasgow, Scotland (Virtual Conference)*, volume 198 of *LIPIcs*, pages 16:1–16:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

5 Markus Anders, Pascal Schweitzer, and Julian Stieß. Engineering a preprocessor for symmetry detection. In *21st International Symposium on Experimental Algorithms, SEA 2023, July 24-26, 2023, Barcelona, Spain*, 2023. to appear.

6 Pascal Schweitzer and Daniel Wiebking. A unifying method for the design of algorithms canonizing combinatorial objects. In Moses Charikar and Edith Cohen, editors, *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, pages 1247–1258. ACM, 2019.

3.8 Thoughts on Computational Design Theory

Brett Stevens (Carleton University – Ottawa, CA)

I will briefly review a history of computational design theory with a bias towards highlighting the breadth of approaches available. I will state three theoretical ideas which seem to be highly relevant to implementing searches: the existence of designs with prescribed automorphism groups, the relationship between designs and codes and the existence of designs derived from other designs. There is a small set of particular algorithms which are heavily used and software systems and implementations which are very useful. I will end with some open problems from large and significant to modest, personal and idiosyncratic.

3.9 Computational methods in finite geometry

Leo Storme (Ghent University, BE)

In finite geometries, many different substructures are investigated. Many are investigated for their geometrical interest, but many are also investigated because of their relevance for other domains, such as coding theory.

Many examples of substructures have not yet been found. It is therefore a good idea to search for examples of substructures in finite geometries. It is also good to classify substructures with computational methods. Some computational methods search specifically for substructures in the respective finite geometries. But there are also many graphs associated to finite geometries, so some searches for substructures can be retranslated to clique or coclique problems in the corresponding graphs.

Concrete problems that can be investigated via computational methods:

- searches for large partial spreads in finite projective spaces and in finite classical polar spaces [4],
- searches for even sets in the projective plane $\mathrm{PG}(2, q)$, q even,
- 2-colorings in Grassmann graphs [3],
- computational searches for open problems on strongly regular graphs [2],
- Cameron-Liebler sets in finite projective spaces and in finite classical polar spaces, and
- the classification of small affine vector subspace partitions in $\mathrm{AG}(5, 2)$ [1].

References

1 J. Bamberg, Y. Filmus, F. Ihringer, and S. Kurz, Affine vector space partitions, preprint, 2022. http://arxiv.org/abs/2211.16561.
2 A.E. Brouwer and H. Van Maldeghem, *Strongly regular graphs.* Encyclopedia of Mathematics and its Applications, 182. Cambridge University Press, Cambridge, 2022.
3 S. De Winter and K. Metsch, Perfect 2-Colorings of the Grassmann Graph of Planes. *Electron. J. Combin.* **27** (2020), no. 1, Paper No. 1.21, 19 pp.
4 S. El-Zanati, H. Jordon, G. Seelinger, P. Sissokho, and L. Spence, The maximum size of a partial 3-spread in a finite vector space over GF(2). *Des. Codes Cryptogr.* **54** (2010), 101-107.

3.10 Open problems on orthogonal(ish) arrays

Ian M. Wanless (Monash University – Clayton, AU)

Joint work of Ian M. Wanless, Michael Gill, Balázs Pozsgay, Marton Mestyan

A famous combinatorial problem is to find the largest set of MOLS (Mutually orthogonal Latin squares) of order 10. After more than 200 years of research on this question, all that is known is that the answer is in the interval [2,6]. In 2014 Dukes and Howard showed that any set of 4 or more MOLS of order 10 must satisfy two non-trivial relations; that is linear dependencies in their incidence matrix. In recent work with my student Michael Gill we have ruled out any relations on pairs of MOLS of order 10. The logical next step is to attempt a computation of triples of MOLS that satisfy two relations. I discussed the feasibility of such a computation and some of the theory of relations.

Another open problem that I discussed was motivated by recent work on multidirectional unitary operators in quantum information theory. That work has inspired us to define a new combinatorial object called a cyclically orthogonal array (COA). This is a variant of traditional orthogonal arrays where we now only require (combinatorial) orthogonality between sets of columns that are cyclically consecutive. Many open problems were discussed, since this is a brand new field. From the physics point-of-view there is interest in COAs that have symmetries which result in space-reflexive or time-reflexive unitary operators.

References

1 M. J. Gill and I. M. Wanless, Pairs of MOLS of order ten satisfying non-trivial relations, *Des. Codes Cryptogr.* **91** (2023), 1293–1313.

2 M. Mestyán, B. Pozsgay and I. M. Wanless, Multi-directional unitarity and maximal entanglement in spatially symmetric quantum states, preprint, 2022. `http://arxiv.org/abs/2210.13017`.

4 Working groups

4.1 Improving reliability and usability of computational projects

Gordon Royle (The University of Western Australia – Crawley, AU)

Joint work of Working group members

We had a wide-ranging discussion on a variety of issues relating to the practice of combinatorial computing.

This included discussions about a number of aspects:

- Making data sets accessible and persistent, even if its creator retires, dies or loses interest.
- Certificates of non-existence of solutions for SAT and/or ILP and CSP solvers.
- What journals should do with respect to computational papers in terms of how much detail of the computation is needed, and how to permit referees and/or readers to replicate the results.
- Discussion of other attempts to tackle this question, such as funding bodies' requirements for "data management plans" and initiatives like MARDI.

- A variety of other anecdotes, personal viewpoints, pointers to successful and unsuccessful examples of how data can or should be shared, discussion on programs such as GAP, Magma and SageMath.

This was all useful and interesting, but it was more a sharing of opinions and knowledge, and we do not have a specific ongoing "problem" to be solved.

4.2 SAT working group

Manfred Scheucher (TU Berlin, DE), Curtis Bright (University of Windsor, CA), Daniel Heinlein (Aalto University, FI), Petteri Kaski (Aalto University, FI), Leonard H. Soicher (Queen Mary University of London, GB), and Alfred Wassermann (Universität Bayreuth, DE)

In this working group we have discussed how to model various mathematical problems. Alfred Wassermann and Leonard Soicher both presented interesting instances of the exact cover problem, but for none of the instances we managed to find a solution or disprove existence yet. Apparently, the clique finding tools by Leonard Soicher still outperformed the sat solvers. Together with Curtis Bright we successfully tackle a problem by Ian Wanless on maximally nonassociative quasigroups. More specifically, while the solver managed to show that there exists a unique example on 9 elements in about 10 CPU hours, further engineering (dynamic symmetry breaking) will be required to progress on 10 and 11 elements.

4.3 Isomorphism Solvers

Pascal Schweitzer (TU Darmstadt, DE)

Among other things, graph isomorphism solvers and canonization tools find application in the context of isomorphism-free exhaustive generation.

The Working Group on Isomorphism Solvers discussed numerous recent developments regarding theoretical and practical aspects of the development of symmetry detection, isomorphism, and canonization tools. The discussions extended to the limits of current solvers, recent new algorithmic ideas, and low-level implementation details in the core subroutines of existing libraries.

Specifically, the following topics were discussed: current implementations of the software libraries NAUTY/TRACES, BLISS, and DEJAVU, the performance of isomorphism solvers on Latin squares and similar combinatorial objects, and efficient implementations of color refinement. We also discussed design questions regarding the interface that isomorphism solvers do, and should, provide. Finally, the current benchmark library for isomorphism testing was discussed. The authors of the symmetry software libraries present in the Working Group expressed that they welcome the challenging instances that other participants may encounter in their research.

4.4 Design theory working group

Ian M. Wanless (Monash University – Clayton, AU), Ilias S. Kotsireas (Wilfrid Laurier University – Waterloo, CA), Denis Krotov (Sobolev Institute of Mathematics – Novosibirsk, RU), and Leo Storme (Ghent University, BE)

A number of open problems in design theory were discussed, involving areas as diverse as Golay pairs, Hadamard matrices, orthogonal arrays and finite geometry.

4.4.1 Cyclically orthogonal arrays

Motivated by recent work in quantum physics we studied objects called cyclically orthogonal arrays (COAs). These are similar to standard orthogonal arrays except that the orthogonality is only required for sets of columns which are consecutive (in a cyclic order). We found that linear cyclic COAs can be built using cyclotomic polynomials. Also linear space- or time-reflexive examples can be made using generator matrices that satisfy particular symmetries. We also found that non-linear COAs can be built from right-inverse-property quasigroups.

4.4.2 Hadamard matrices

A conjecture of O'Cathain and Wanless [1] states that any trade in a complex Hadamard matrix of order n must contain at least n entries. This conjecture is known to be true for real Hadamard matrices, and for trades that result from multiplying a rectangular subarray by a scalar. A special case of Hadamard matrices with a trade of size n placed in diagonal entries has a skew structure in the real case and skew or mixed-skew structure in the case of complex Hadamard $\{1, -1, i, -i\}$-matrices. The existence of mixed-skew Hadamard matrices is known only for several small orders. Bicyclic mixed-skew complex Hadamard matrices of order $2n$ are equivalent to i-reversible periodic complex Golay pairs of sequences of length n, whose existence is also unknown for large n.

4.4.3 Golay pairs

Periodic Golay pairs constitute the periodic analog of the well-studied Golay pairs. The computational state-of-the-art in Periodic Golay pairs is contained in the 3 papers [2, 3, 4]. We pose as an open problem the construction of Periodic Golay pairs of order 90. Given that 90 is a highly composite number, the method of compression seems like an appropriate tool.

4.4.4 Finite geometry

See the separate abstract by Leo Storme in Section 3.9.

References

1 P. Ó Catháin and I. M. Wanless, Trades in complex Hadamard matrices, in C. J. Colbourn (ed.), *Algebraic Design Theory and Hadamard Matrices*, Springer Proceedings in Mathematics and Statistics **133** (2015), 213–221.
2 D.Z. Djoković, I.S. Kotsireas, Periodic Golay pairs of length 72. Algebraic design theory and Hadamard matrices, 83–92, Springer Proc. Math. Stat., 133, Springer, Cham, 2015.
3 D.Z. Djoković, I.S. Kotsireas, Some new periodic Golay pairs. Numer. Algorithms 69 (2015), no. 3, 523–530.
4 D.Z. Djoković, I.S. Kotsireas, D. Recoskie, J. Sawada, Charm bracelets and their application to the construction of periodic Golay pairs. Discrete Appl. Math. 188 (2015), 32–40.

4.5 Tactical decompositions working group

Alfred Wassermann (Universität Bayreuth, DE)

Combinatorial designs. A combinatorial t-(v, k, λ) design $(V, \mathcal{D})$ is a set V consisting of v *points* together with a set $\mathcal{D}$ of k-subsets of V called *blocks* such that each t-subset of V is contained in exactly λ blocks, see e.g. [1, 2] or [7]. The q-analogs of combinatorial designs are called subspace designs, see [5] for an introduction and overview.

It is well known that a t-(v, k, λ) design $(V, \mathcal{D})$ is also an s-(v, k, λ_s) design for $0 \le s \le t$ where

$$\lambda_s = \lambda \frac{\binom{v-s}{t-s}}{\binom{k-s}{t-s}} .$$

In particular, λ_0 is the number of blocks of the design and λ_1 is the number of blocks each point is contained in, which is called the *replication number*. It is also well known that for a t-(v, k, λ) design the number of blocks which contain a given i-set of points and are disjoint to a given j-set of points is equal to

$$\lambda_{i,j} = \lambda \frac{\binom{v-i-j}{k-j}}{\binom{v-t}{k-t}} ,$$

see e.g. [7, II.4.2, p. 80].

Incidence matrices. The $v \times \lambda_0$ point-block incidence matrix N of a t-(v, k, λ) design $(V, \mathcal{D})$ is defined by

$$N_{P,B} = \begin{cases} 1, & \text{if } P \in B, \\ 0, & \text{otherwise} \end{cases}$$

for $P \in V$ and $B \in \mathcal{D}$. Bose [4] showed for the point-block incidence matrix N of a 2-(v, k, λ) design the equation

$$NN^\top = \lambda_1 I + \lambda(J - I), \tag{1}$$

where I is the $v \times v$ identity matrix and J is the $v \times v$ all-ones matrix.

For a general t-(v, k, λ) design with $t \ge 2$, the $\binom{v}{e} \times \lambda_0$ higher incidence matrix $N^{(e)}$ for $e \le k$ is defined by

$$N^{(e)}_{E,B} = \begin{cases} 1, & \text{if } E \subset B, \\ 0, & \text{otherwise} \end{cases}$$

for $E \in \binom{V}{e}$ and $B \in \mathcal{D}$. The $\binom{v}{s} \times \binom{v}{e}$ incidence matrix $W^{(se)}$ between s-subsets and *all* e-subsets of V is defined by

$$W^{(se)}_{S,E} = \begin{cases} 1, & \text{if } S \subset E, \\ 0, & \text{otherwise} \end{cases}$$

for $S \in \binom{V}{s}$ and $E \in \binom{V}{e}$. Wilson [18] showed for $e + f \leq t$ the equation

$$N^{(e)}\,(N^{(f)})^\top = \sum_{i=0}^{\min\{e,f\}} \lambda_{e+f-i,\,i}(W^{(ie)})^\top\,W^{(if)}\,. \tag{2}$$

Note that $N^{(e)}(N^{(f)})^\top$ contains in the row labeled by the e-subset E and in the column labeled by the f-subset F the number of blocks of the design which contain both E and F. It is clear that this number is $\lambda_{e+f-\mu}$ with $\mu = \#(E \cap F)$, i. e.

$$\left(N^{(e)}\,(N^{(f)})^\top\right)_{E,F} = \lambda_{\#(E\cup F)}\,.$$

Also in [18], Wilson proved among others the equation

$$W^{(ie)}\,N^{(e)} = \binom{k-i}{e-i}\,N^{(i)} \qquad \text{for } 0 \leq i \leq e \leq k\,. \tag{3}$$

For $e = 1$ and $i = 0$ equation (3) simply states that each block of the design contains k points.

Tactical decomposition matrices. The use of *tactical decompositions* in design theory has been initiated by Dembowski [9], see also [10] and Beutelspacher [3, pp. 210–220]. Dembowski's main interest was to use tactical decompositions to study properties of symmetric designs. From an algorithmic point of view, tactical decompositions were first used by Janko and Tran Van Trung [11] to construct symmetric $(78, 22, 6)$ designs. Their method was picked up and generalized in numerous papers, see [6, 8, 14] to name just a few. In [17] the use of tactical decompositions has been generalized to subspace designs.

Dembowski [9, 10] studied tactical decompositions of incidence structures from group actions, see also Beutelspacher [3, pp. 210–220].

Let $(V, \mathcal{D})$ be a 2-(v, k, λ) design invariant under some group G. The action of G partitions V into orbits $\mathcal{P}_1, \ldots, \mathcal{P}_m$ and $\mathcal{D}$ into orbits $\mathcal{B}_1, \ldots, \mathcal{B}_n$. Let N be the point-block incidence matrix of $(V, \mathcal{D})$ and for $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ let $N_{i,j}$ be the submatrix of N whose rows are assigned to the elements $\mathcal{P}_i$ and whose columns to the elements of $\mathcal{B}_j$. Then $N_{i,j}$ has a constant number of ones in each row and a constant number of ones in each column. Such a decomposition of N into submatrices $N_{i,j}$ is called *tactical*.

If we replace for all i, j the submatrix $N_{i,j}$ by the number of ones in each row we get an $(m \times n)$-matrix ρ, and if we replace the submatrix $N_{i,j}$ by the number of ones in each column we get an $(m \times n)$-matrix κ. The matrices ρ and κ are both called *tactical decomposition matrix*. In [9] the following properties of ρ and κ and the matrices $P = \operatorname{diag}(\#\mathcal{P}_i)$ and $B = \operatorname{diag}(\#\mathcal{B}_i)$ are shown:

$$P \cdot \rho = \kappa \cdot B \tag{4}$$

$$\rho \cdot (1, \ldots, 1)^\top = (\lambda_1, \ldots, \lambda_1)^\top \tag{5}$$

$$(1, \ldots, 1) \cdot \kappa = (k, \ldots, k) \tag{6}$$

$$\rho \cdot \kappa^\top = (\lambda_1 - \lambda) \cdot I + \lambda \cdot P \cdot J \tag{7}$$

The general approach outlined by Janko and Tran Van Trung is to first enumerate all tactical decomposition matrices of designs with prescribed automorphisms up to permutations of rows and columns. For this, Dembowski [9] has given powerful constraints for a matrix to be a tactical decomposition of the point-block incidence matrix of a 2-design. In a second step, all remaining tactical decomposition matrices are expanded – if possible – to point-block incidence matrices of designs.

Compared with the well-known method of Kramer and Mesner [13] which also restricts the search space to designs with prescribed automorphisms, the method of Janko and Tran Van Trung has the advantage that it is not necessary to compute all orbits of k-subsets of V and therefore allows the search for 2-(v,k,λ) designs with larger k and smaller automorphism group. The drawback however is that it does not reduce the search space if the prescribed group of automorphisms is point-transitive, and that it seemed to be restricted to 2-designs for a long time.

In recent years two generalizations to t-designs for $t \geq 2$ have been published.

The first generalization in [14, 15, 16] gives constraints for the tactical decomposition of the point-block incidence matrix of a t-design of general strength $t \geq 2$:

▶ **Theorem 1.** *Let $1 \leq s \leq t$ and $m_1, \ldots, M_s$ be positive integers, such that $m_s + \cdots + m_s \leq t$. Let $\mathfrak{P}_{i_1}, \ldots, \mathfrak{P}_{i_s}$ be mutually distinct. Then*

$$\sum_{j=1}^{n} \rho_{i_1 j} \kappa_{i_1 j}^{m_1-1} \kappa_{i_2 j}^{m_2} \cdots \kappa_{i_s j}^{m_s} = \sum_{\omega \in \Omega} \lambda_{\omega_1+\cdots+\omega_s} \left\{ {m_1 \atop \omega_1} \right\} (\#\mathfrak{P}_{i_1} - -1)_{\omega_1 - 1} \prod_{j=2}^{s} \left\{ {m_j \atop \omega_j} \right\} (\#\mathfrak{P}_{i_j})_{\omega_j},$$

where

$$\Omega = \{(\omega_1, \ldots, \omega_s) \mid 1 \leq \omega_j \leq m_i\},$$

$\left\{ {n \atop k} \right\}$ *are the Stirling numbers of second kind and $(x)_n = \prod_{i=0}^{n-1}(x-i)$ is the falling factorial.*

The second generalization in [12] defines higher tactical decomposition matrices and shows the Wilson's equations can be generalized to these higher tactical decomposition matrices.

For $x \in \{0, \ldots, v\}$, let $\mathfrak{P}_x$ be a partition of the set $\binom{V}{x}$. The part of $\mathfrak{P}_x$ containing some $X \in \binom{V}{x}$ will be denoted by $[X]$. We call $(\mathfrak{P}_0, \ldots, \mathfrak{P}_v)$ a *tactical sequence of partitions* on V if for all $x, y \in \{0, \ldots, v\}$ with $x \leq y$ and for all $[X], \mathcal{X} \in \mathfrak{P}_x$ and $[Y], \mathcal{Y} \in \mathfrak{P}_y$, the numbers

$$R^{(xy)}_{[X],\mathcal{Y}} = \#\{Y \in \mathcal{Y} \mid X \subseteq Y\} \qquad \text{and} \qquad K^{(xy)}_{\mathcal{X},[Y]} = \#\{X \in \mathcal{X} \mid X \subseteq Y\}$$

are well-defined, i. e. they do not depend on the choice of the representative X of $[X]$ nor of the representative Y of $[Y]$. In this case, the above defined numbers yield matrices $R^{(xy)}, K^{(xy)} \in \mathbb{Z}^{\mathfrak{P}_x \times \mathfrak{P}_y}$. A common source of tactical sequences of partitions are permutation groups $G \leq S_V$, where for all $x \in \{0, \ldots, v\}$ the partition $\mathfrak{P}_x$ is the set of orbits of the induced action of G on $\binom{V}{x}$.

In the following, a tactical sequence $(\mathfrak{P}_0, \ldots, \mathfrak{P}_v)$ of partitions on V is fixed. The matrices $R^{(xx)}$ and $K^{(xx)}$ are $\#\mathfrak{P}_x \times \#\mathfrak{P}_x$ identity matrices. The matrices $R^{(0x)}$ and $K^{(0x)}$ are of size $1 \times \#\mathfrak{P}_x$, where all entries of $K^{(0x)}$ are 1, and $R^{(0x)}$ contains the part sizes, i. e. $R^{(0x)}_{\{\emptyset\},\mathcal{X}} = \#\mathcal{X}$.

Let $(V, \mathcal{D})$ be t-(v,k,λ) design with $t \leq k \leq v-t$, such that the block set is the union of parts in $\mathfrak{P}_k$, i. e. $\mathcal{D} = \bigcup \mathfrak{B}$ with $\mathfrak{B} \subseteq \mathfrak{P}_k$.

For $x \in \{0, \ldots, k\}$ we define the tactical decomposition matrices $\rho^{(x)}, \kappa^{(x)} \in \mathbb{Z}^{\mathfrak{P}_x \times \mathfrak{B}}$ via

$$\rho^{(x)}_{[X],\mathcal{B}} = \#\{B \in \mathcal{B} \mid X \subseteq B\} \qquad \text{and} \qquad \kappa^{(x)}_{\mathcal{X},[B]} = \#\{X \in \mathcal{X} \mid X \subseteq B\}.$$

By the properties of the fixed tactical sequence $(\mathfrak{P}_0, \ldots, \mathfrak{P}_v)$ of partitions on V, this definition does not depend on the choice of the representatives. Note that $\rho^{(x)}$ is the restriction of $R^{(xk)}$ to the columns whose labels are contained in $\mathfrak{B}$. In particular, $\rho^{(0)}$ and $\kappa^{(0)}$ are of size $1 \times \#\mathfrak{B}$, where all entries of $\kappa^{(0)}$ are 1, and $\rho^{(0)}$ contains the sizes of the block parts.

In [12] the following theorem has been proved for these higher tactical decomposition matrices.

▶ **Theorem 2.** *Let V be a finite set of size v and let $(\mathfrak{P}_0, \ldots, \mathfrak{P}_v)$ be a tactical sequence of partitions on V. Let $(V, \mathcal{D})$ be a non-empty t-(v, k, λ) design with $t \le k \le v - t$, such that the block set has the form $\mathcal{D} = \bigcup \mathfrak{B}$ with $\mathfrak{B} \subseteq \mathfrak{P}_k$. Let e, f be non-negative integers with $e + f \le t$.*

Then

$$\rho^{(e)} \, (\kappa^{(f)})^\top = \sum_{j=0}^{\min(e,f)} \lambda_{e+f--j,\, j} \, (K^{(je)})^\top R^{(jf)}.$$

Moreover, for non-negative integers x, y with $x \le y \le k$

$$R^{(xy)} \, \rho^{(y)} = \binom{k-x}{y-x} \rho^{(x)} \qquad \text{and} \qquad K^{(xy)} \, \kappa^{(y)} = \binom{k-x}{y-x} \kappa^{(x)}.$$

It should be noted that the above theorem has a q-analog version for subspace designs.

Summary of the work. In the working group an introduction to tactical decomposition matrices has been given. This was followed by a long discussion in which the two generalizations were compared and where it was attempted to derive the first generalization from the second. As a result, the exact relation of the two generalizations could not be determined easily and will be subject of future research work.

Next, algorithmic use of the second generalization was studied. The participants of the working group agreed that the approach by Janko and Tran van Trung might be generalized to higher tactical decompositions, however its applicability will be restricted to a very specific set of parameters of combinatorial designs and choice of prescribed automorphism groups.

As a followup, research groups from Zagreb and Bayreuth are planning to write a proposal for a joint research project to develop a practical algorithm based on higher tactical decomposition matrices.

Another topic of the working group was the indexing step of the Janko-Tran van Trung algorithm and possible generalizations.

Finally, for graphs and hypergraphs the concept of tactical decomposition matrices is known as *equitable partitions*. Main contributions for these have been given by the research group in Novosibirsk. The presentation of results on equitable partitions lead to surprising connections to the so called Hartman's conjecture in design theory. This was explored and further research on this connection seems to be promising.

References

1. T. Beth, D. Jungnickel, and H. Lenz. *Design theory*. 2nd ed. Vol. I. Encyclopedia of Mathematics and its Applications 69. Cambridge: Cambridge University Press, 1999. DOI: 10.1017/CBO9780511549533.
2. T. Beth, D. Jungnickel, and H. Lenz. *Design theory*. 2nd ed. Vol. II. Encyclopedia of Mathematics and its Applications 78. Cambridge: Cambridge University Press, 1999. DOI: 10.1017/CBO9781139507660.
3. A. Beutelspacher. *Einführung in die endliche Geometrie*. Vol. I: *Blockpläne*. Mannheim: Bibliographisches Institut, 1982.
4. R. C. Bose. "A note on Fisher's inequality for balanced incomplete block designs". In: *Ann. Math. Statistics* 20.4 (1949), pp. 619–620. DOI: 10.1214/aoms/1177729958.
5. M. Braun, M. Kiermaier, and A. Wassermann. "q-analogs of designs: subspace designs". In: *Network coding and subspace designs*. Ed. by M. Greferath, M. O. Pavčević, N. Silberstein, and M. Á. Vázquez-Castro. Signals Commun. Technol. Cham: Springer, 2018, pp. 171–211. DOI: 10.1007/978-3-319-70293-3_8.

6 V. Ćepulić. "On symmetric block designs $(40, 13, 4)$ with automorphisms of order 5". In: *Discrete Math.* 128.1 (1994), pp. 45–60. DOI: 10.1016/0012-365X(94)90103-1.

7 C. J. Colbourn and J. H. Dinitz, eds. *Handbook of combinatorial designs*. 2nd ed. Discrete Mathematics and its Applications. Boca Raton, FL: Chapman & Hall/CRC, 2007. DOI: 10.1201/9781420010541.

8 D. Crnković and S. Rukavina. "Construction of block designs admitting an Abelian automorphism group". In: *Metrika* 62.2–3 (2005), pp. 175–183. DOI: 10.1007/s00184-005-0407-y.

9 P. Dembowski. "Verallgemeinerungen von Transitivitätsklassen endlicher projektiver Ebenen". In: *Math. Z.* 69.1 (1958), pp. 59–89. DOI: 10.1007/BF01187393.

10 P. Dembowski. *Finite geometries*. Ergebnisse der Mathematik und ihrer Grenzgebiete 44. Berlin and New York: Springer-Verlag, 1968.

11 Z. Janko and T. V. Trung. "Construction of a new symmetric block design for $(78, 22, 6)$ with the help of tactical decompositions". In: *J. Combin. Theory, Ser. A* 40.2 (1985), pp. 451–455. DOI: 10.1016/0097-3165(85)90107-4.

12 M. Kiermaier and A. Wassermann. "Higher incidence matrices and tactical decomposition matrices". In: *Glasnik Matematički* to appear (2023).

13 E. S. Kramer and D. M. Mesner. "t-designs on hypergraphs". In: *Discrete Math.* 15.3 (1976), pp. 263–296. DOI: 10.1016/0012-365X(76)90030-3.

14 V. Krčadinac, A. Nakić, and M. O. Pavčević. "The Kramer-Mesner method with tactical decompositions: some new unitals on 65 points". In: *J. Combin. Des.* 19.4 (2011), pp. 290–303. DOI: 10.1002/jcd.20277.

15 V. Krčadinac, A. Nakić, and M. O. Pavčević. "Equations for coefficients of tactical decomposition matrices for t-designs". In: *Des. Codes Cryptogr.* 72.2 (2014), pp. 465–469. DOI: 10.1007/s10623-012-9779-y.

16 A. Nakić. "Non-existence of a simple 3-$(16, 7, 5)$ design with an automorphism of order 3". In: *Discrete Math.* 338.4 (2015), pp. 555–565. DOI: 10.1016/j.disc.2014.11.016.

17 A. Nakić and M. O. Pavčević. "Tactical decompositions of designs over finite fields". In: *Des. Codes Cryptogr.* 77.1 (2015), pp. 49–60. DOI: 10.1007/s10623-014-9988-7.

18 R. M. Wilson. "Incidence matrices of t-designs". In: *Linear Algebra Appl.* 46 (1982), pp. 73–82. DOI: 10.1016/0024-3795(82)90027-1.

Participants

- Thais Bardini Idalino
Federal University of Santa Catarina, BR
- Curtis Bright
University of Windsor, CA
- Jan De Beule
Free University of Brussels, BE
- Alena Ernst
Universität Paderborn, DE
- Tuvi Etzion
Technion – Haifa, IL
- Marcus Greferath
University College Dublin, IE
- Daniel Heinlein
Aalto University, FI
- Delaram Kahrobaei
City University of New York, US
- Petteri Kaski
Aalto University, FI
- Sandra Kiefer
University of Oxford, GB
- Michael Kiermaier
Universität Bayreuth, DE
- Ilias S. Kotsireas
Wilfrid Laurier University – Waterloo, CA
- Vedran Krcadinac
University of Zagreb, HR
- Aleksandr Krotov
Innopolis University, RU
- Denis Krotov
Sobolev Institute of Mathematics – Novosibirsk, RU
- Sascha Kurz
Universität Bayreuth, DE
- Pekka Lampio
Aalto University, FI
- Brendan McKay
Australian National University – Acton, AU
- Lucia Moura
University of Ottawa, CA
- Gábor P. Nagy
Budapest University of Technology & Economics, HU
- Anamari Nakic
University of Zagreb, HR
- Eamonn O'Brien
University of Auckland, NZ
- Patric Östergård
Aalto University, FI
- Mario Osvin Pavcevic
University of Zagreb, HR
- David Pike
Memorial University of Newfoundland – St. John's, CA
- Adolfo Piperno
Sapienza University of Rome, IT
- Gordon Royle
The University of Western Australia – Crawley, AU
- Manfred Scheucher
TU Berlin, DE
- Pascal Schweitzer
TU Darmstadt, DE
- Leonard H. Soicher
Queen Mary University of London, GB
- Brett Stevens
Carleton University – Ottawa, CA
- Leo Storme
Ghent University, BE
- Ana Sumberac
University of Rijeka, HR
- Edwin van Dam
Tilburg University, NL
- Ian M. Wanless
Monash University – Clayton, AU
- Alfred Wassermann
Universität Bayreuth, DE
- Charlene Weiß
Universität Paderborn, DE

Report from Dagstuhl Seminar 23162

New Frontiers of Parameterized Complexity in Graph Drawing

Robert Ganian[*1], **Fabrizio Montecchiani**[*2], **Martin Nöllenburg**[*3], **Meirav Zehavi**[*4], **and Liana Khazaliya**[†5]

1 **TU Wien, AT.** `rganian@gmail.com`
2 **University of Perugia, IT.** `fabrizio.montecchiani@unipg.it`
3 **TU Wien, AT.** `noellenburg@ac.tuwien.ac.at`
4 **Ben Gurion University – Beer Sheva, IL.** `zehavimeirav@gmail.com`
5 **TU Wien, AT.** `liana.khazaliya@gmail.com`

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 23162 "New Frontiers of Parameterized Complexity in Graph Drawing". The seminar was held in-person from April 16 to April 21, 2023. It brought together 32 researchers from the Graph Drawing and the Parameterized Complexity research communities to discuss and explore new research frontiers on the interface between the two fields. The report collects the abstracts of talks and open problems presented in the seminar, as well as brief progress reports from the working groups.

Seminar April 16–21, 2023 – https://www.dagstuhl.de/23162
2012 ACM Subject Classification Theory of computation → Parameterized complexity and exact algorithms; Theory of computation → Computational geometry
Keywords and phrases algorithm design, computational geometry, graph drawing, parameterized complexity
Digital Object Identifier 10.4230/DagRep.13.4.58

1 Executive Summary

Robert Ganian (TU Wien, AT)
Fabrizio Montecchiani (University of Perugia, IT)
Martin Nöllenburg (TU Wien, AT)
Meirav Zehavi (Ben Gurion University – Beer Sheva, IL)

In modern life, it is of paramount importance that computational tasks be performed both precisely and efficiently. Thus, the design and analysis of algorithms to execute such tasks lie at the heart of computer science. However, since the proof of the Cook-Levin theorem in the 1970s, numerous problems have been shown to be NP-hard. Fortunately, the field of *parameterized complexity* [7, 5, 10, 20], initiated in the 1980s by Downey and Fellows, yields (in)tractability results that are both deeper and fine-grained. Specifically, it shows that the "hardness" of an NP-hard problem can often be traced to particular *parameters* of its instances.

* Editor / Organizer
† Editorial Assistant / Collector

In a nutshell, a *parameterization* of a problem Π is the association of a parameter k with every instance of Π, capturing how "structured" the instance is. The main goal in parameterized complexity is to confine the combinatorial explosion in the runtime to a suitable parameter k, rather than to let it depend on the entire input size n. So, fixed-parameter algorithms naturally "scale" with the amount of structure of a given instance. Formally, a problem is *fixed-parameter tractable* (in short, in FPT) with respect to a parameter k if it can be solved by an algorithm, called a *fixed-parameter algorithm*, whose runtime is bounded by $f(k) \cdot n^{O(1)}$ for some computational function f of k. Over the past four decades, the parameterized complexity paradigm has yielded a rich and deep theory, with powerful approaches to obtain fixed-parameter algorithms for a variety of problems as well as machinery that can rule out such algorithms. For problems in FPT, it also offers the necessary tools to develop the fastest possible fixed-parameter algorithms (often with tight conditional lower bounds).

While the parameterized paradigm can be applied in a wide range of different fields, the focus of this Dagstuhl Seminar lay squarely on its potential synergies with *graph drawing*, a self-standing discipline that has evolved tremendously over the last decades. Indeed, given the ubiquity of graphs in many fields of science and technology, there is a strong interest in algorithms that can provide effective graphical representations of graphs, for the sake of both analysis and communication. Today, graph drawing is a mature area of computer science [6, 17, 21, 22] with its own annual conference, the International Symposium on Graph Drawing and Network Visualization (GD)[1]. The focus of graph drawing as a research area today is on both fundamental and practical aspects, such as combinatorics and algorithm design on the one hand, and the development of network visualization systems and interfaces on the other hand. At a very high level, graph drawing deals with the construction and analysis of geometric representations of graphs and networks subject to specific layout conventions and constraints, such as different notions of planarity or more general crossing constraints, grid layouts, orthogonal drawings, and many more. Notably, this gives rise to many computational problems that are NP-hard in the classical sense but naturally multivariate, making parameterized analysis particularly attractive. Yet, so far, research at the intersection of parameterized complexity and graph drawing has not reached its full potential.

Parameterized Graph Drawing. Most of the early efforts to conduct research at the intersection of parameterized complexity and graph drawing have been directed at variants of the classic Crossing Minimization problem, introduced by Turán in 1940 [23], parameterized by the number of crossings. Here, the objective is to draw a given graph in the plane so as to induce the minimum number of crossings, which was shown to be FPT already in 2001 [13]. A few subsequent works followed [18, 15], including the best paper of GD 2019 [16]. Other examples of graph drawing problems successfully targeted by the parameterized paradigm include crossing minimization in restricted settings [8, 19], crossing-sensitive subgraph detection [1, 14], parameterized algorithms for stack and queue layouts [2, 3, 4], and drawing extension problems [9, 11]. While these and other works already lay down the foundations of possible positive synergies between the two fields, there still is a multitude of problems and opportunities in graph drawing where parameterized analysis would be beneficial and has not yet been carried out.

[1] see `www.graphdrawing.org`

Seminar Goals

The main goal of the seminar was to explore and initiate collaborations at the intersection of Graph Drawing and Parameterized Complexity, with emphasis on new research frontiers. In particular:

- First, the seminar focused on prominent topics in Graph Drawing, encouraging the consideration of topics that have been only little – or not at all – studied from the perspective of Parameterized Complexity. Here, the discussions centered on the identification of open problems of wide interest as well as directions for future research.
- Second, the seminar focused on new tools and sub-areas of Parameterized Complexity, such as structural parameterization, non-standard general decomposition theorems, and FPT-approximation, which have been rarely – or not at all – examined in the context of Graph Drawing. The aim was to understand their potential and relevance in the context of the aforementioned topics.

The seminar included invited talks focused on the above goals, setting the ground for the participants to propose problems (and, more generally, research directions) to work on, as well as be familiar with relevant tools to work with. The final selection of problems targeted by working groups was carried out during the seminar itself. This, in turn, also resulted in the establishment of new and fruitful collaborations.

Seminar Program

We started the seminar week on Monday morning with short self-introduction presentations of all participants, followed by the four invited overview talks mentioned above, which aimed to set the grounds for the research discussions throughout the week. We invited two speakers from each of the two communities coming together in this seminar to create a common understanding of research questions, tools, and terminology. The first of these lectures was given by Saket Saurabh on the topic of fixed-parameter tractability for geometric intersection graphs. Next, Ignaz Rutter spoke about the widely used SPQR tree data structure for representing planar graph embeddings. As the third speaker, Bart M. P. Jansen highlighted three frontiers in parameterized graph drawing based on lossy (structural) kernelization and hybrid parameterizations. Lastly, Giordano Da Lozzo presented complexity results and parameterized algorithms for the upward book embedding problem. Further, we had a short session for reports about the final results of working groups from the previous edition of the seminar in 2021. Section 3 contains more detailed abstracts of the invited lectures and the reports from Dagstuhl Seminar 21293.

The remainder of the seminar focused on research discussions on new frontiers of parameterized graph drawing. In order to identify these frontiers, we had two open problem sessions, one on Monday afternoon and one on Tuesday morning. A total of 13 open research questions (see Section 4) were contributed by the participants. Based on a preference vote, we created five working groups for the seminar week, each having six or seven participants mixed from the two communities joined in this seminar. The following five topics were investigated in depth:

1. Parameterized Complexity of the Maximum Bimodal Subgraph Problem
2. Parameterized Complexity of Upward Planarity and st-Planar Edge Completion Problems
3. Parameterized Algorithms for the Sequential Crossing Number Problem
4. Deleting Few Edges from Ordered Graphs to Make them k-Plane
5. Clustered Planarity and Planar $\mathcal{F}$-Augmentations

For each of the five working groups, Section 5 contains a detailed report of their current state of research. During the seminar, on Tuesday and Thursday afternoon, we had plenary sessions for the groups' progress reports sharing their latest results and further plans. Wednesday afternoon was reserved for the traditional excursion, for which we enjoyed a Treetop-Walk near the Great Bend in the Saar and then hiked to Mettlach for a traditional brewery dinner.

Future Plans

The composition of the working groups was made having in mind two main criteria: balancing the number of participants from the two communities, guaranteeing the presence of some young researcher in each group. This strategy was indeed successful and it led to new collaborations between researchers in the graph drawing and parameterized complexity communities, as well as to new opportunities for young researchers. As a primary outcome of the seminar, we expect research papers published at the core conferences and journals for the graph drawing and parameterized complexity communities, for instance:

- The International Symposium on Computational Geometry (SoCG),
- The International Symposium on Graph Drawing and Network Visualization (GD),
- The ACM-SIAM Symposium on Discrete Algorithms (SODA), and
- The International Symposium on Algorithms and Computation (ISAAC).

In the mid- and long-term horizon, the seminar will also help to gradually blend the boundary between the two communities, with young researchers building their academic journey on the intersection of the two fields. In this respect, the seminar can also lead to the development of new parameterized tools and techniques that are designed to deal with the specific obstacles that arise when trying to apply parameterized approaches in the graph drawing setting.

This Dagstuhl Seminar and its predecessor (Dagstuhl Seminar 21293) have revealed, in a systematic way, the astounding wealth of problems in Graph Drawing that are naturally multivariate and hence suitable for parameterized analysis. The two communities are now aware of these problems, and the above mentioned conferences (among others) publish every year a growing number of results making progress on these problems. Besides this, informal feedback from the communities confirms the great interest in having such seminars on a more regular basis. We therefore plan to have follow-up seminars to support the growth of parameterized complexity in graph drawing.

Evaluation

Taking into account the results of the Dagstuhl survey conducted after the seminar and informal feedback to the organizers, it is fair to say that the seminar was highly appreciated. The participants provided significantly above-average scores for the seminar, particularly in areas concerning new collaborative research projects, identification of novel research directions, and collaboration between different research communities. Overall, we believe that the seminar's goals of identifying new research directions and initiating collaborations at the intersection of the two different fields of Graph Drawing and Parameterized Complexity were very successful. We are looking forward to seeing the first scientific outcomes of the seminar in the near future and to continuing the efforts to support the growth of interest in parameterized analysis of problems in Graph Drawing.

Acknowledgments

Schloss Dagstuhl, once again, was a perfect location for hosting this research seminar with its great conference and meeting facilities, combined with the unique atmosphere of the castle and lots of opportunities for socializing and networking. On behalf of all participants, we want to express our gratitude to the entire Dagstuhl staff for letting us have such a wonderful week. We further thank Liana Khazaliya for helping us collect the various contributions and prepare this report.

References

1. Akanksha Agrawal, Grzegorz Guspiel, Jayakrishnan Madathil, Saket Saurabh, and Meirav Zehavi. Connecting the Dots (with Minimum Crossings). In Gill Barequet and Yusu Wang, editors, *Symposium on Computational Geometry (SoCG 2019)*, volume 129 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 7:1–7:17, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
2. Michael J. Bannister, David Eppstein, and Joseph A. Simons. Fixed parameter tractability of crossing minimization of almost-trees. In *Graph Drawing (GD 2013)*, volume 8242 of *Lecture Notes in Computer Science*, pages 340–351. Springer, 2013.
3. Sujoy Bhore, Robert Ganian, Fabrizio Montecchiani, and Martin Nöllenburg. Parameterized algorithms for book embedding problems. *J. Graph Algorithms Appl.*, 24(4):603–620, 2020.
4. Sujoy Bhore, Robert Ganian, Fabrizio Montecchiani, and Martin Nöllenburg. Parameterized algorithms for queue layouts. In *Graph Drawing*, volume 12590 of *Lecture Notes in Computer Science*, pages 40–54. Springer, 2020.
5. Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms.* Springer, 2015.
6. Giuseppe Di Battista, Peter Eades, Roberto Tamassia, and Ioannis G. Tollis. *Graph Drawing: Algorithms for the Visualization of Graphs.* Prentice-Hall, 1999.
7. Rodney G. Downey and Michael R. Fellows. *Fundamentals of Parameterized Complexity.* Texts in Computer Science. Springer Verlag, 2013.
8. Vida Dujmovic, Michael R. Fellows, Matthew Kitching, Giuseppe Liotta, Catherine McCartin, Naomi Nishimura, Prabhakar Ragde, Frances A. Rosamond, Sue Whitesides, and David R. Wood. On the parameterized complexity of layered graph drawing. *Algorithmica*, 52(2):267–292, 2008.
9. Eduard Eiben, Robert Ganian, Thekla Hamm, Fabian Klute, and Martin Nöllenburg. Extending partial 1-planar drawings. In *ICALP*, volume 168 of *LIPIcs*, pages 43:1–43:19. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020.
10. F. V. Fomin, D. Lokshtanov, S. Saurabh, and M. Zehavi. *Kernelization: Theory of Parameterized Preprocessing.* Cambridge University Press, 2018.
11. Robert Ganian, Thekla Hamm, Fabian Klute, Irene Parada, and Birgit Vogtenhuber. Crossing-optimal extension of simple drawings. In *ICALP*, volume 198 of *LIPIcs*, pages 72:1–72:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021.
12. Ashim Garg and Roberto Tamassia. On the computational complexity of upward and rectilinear planarity testing. *SIAM J. Comput.*, 31(2):601–625, 2001.
13. Martin Grohe. Computing crossing numbers in quadratic time. *J. Comput. Syst. Sci.*, 68(2):285–302, 2004.
14. Magnús M. Halldórsson, Christian Knauer, Andreas Spillner, and Takeshi Tokuyama. Fixed-parameter tractability for non-crossing spanning trees. In *Algorithms and Data Structures (WADS 2007)*, volume 4619 of *Lecture Notes in Computer Science*, pages 410–421. Springer, 2007.

15 Petr Hlinený and Marek Dernár. Crossing number is hard for kernelization. In *Symposium on Computational Geometry (SoCG 2016)*, volume 51 of *LIPIcs*, pages 42:1–42:10. Schloss Dagstuhl – Leibniz-Zentrum fuer Informatik, 2016.
16 Petr Hlinený and Abhisekh Sankaran. Exact crossing number parameterized by vertex cover. In *Graph Drawing and Network Visualization (GD 2019)*, volume 11904 of *Lecture Notes in Computer Science*, pages 307–319. Springer, 2019.
17 Michael Jünger and Petra Mutzel, editors. *Graph Drawing Software*. Springer, 2004.
18 Ken-ichi Kawarabayashi and Bruce A. Reed. Computing crossing number in linear time. In *Symposium on Theory of Computing (STOC 2007)*, pages 382–390. ACM, 2007.
19 Fabian Klute and Martin Nöllenburg. Minimizing crossings in constrained two-sided circular graph layouts. *J. Comput. Geom.*, 10(2):45–69, 2019.
20 Rolf Niedermeier. *Invitation to Fixed-Parameter Algorithms*. Oxford Lecture Series in Mathematics and Its Applications. OUP Oxford, 2006.
21 Takao Nishizeki and Md. Saidur Rahman. *Planar Graph Drawing*, volume 12 of *Lecture Notes Series on Computing*. World Scientific, 2004.
22 Roberto Tamassia, editor. *Handbook on Graph Drawing and Visualization*. Chapman and Hall/CRC, 2013.
23 Paul Turán. A note of welcome. *Journal of Graph Theory*, 1(1):7–9, 1977.

2 Table of Contents

3 Overview of Talks

3.1 The Hitchhiker's Guide to SPQR-trees

Ignaz Rutter (Universität Passau, DE)

SPQR-trees are a widely used tool in Graph Drawing that is at the heart of a large body of algorithmic results that are concerned with embeddings of biconnected planar graphs. Their key property is that they break down the complicated task of choosing a planar embedding of a biconnected graph into multiple independent choices, each of which is either binary or consists in arbitrarily choosing a permutation of a set of parallel subgraphs between two pole vertices.

The talk gives an introduction to SPQR-trees that is aimed at a broad audience. To cater to different backgrounds, we first give three different definitions of SPQR-trees that are complementary to each other with respect to the properties that can be easily deduced from them. In the second part of the talk we illustrate how to use SPQR-trees to solve constrained embedding problems, i.e., how to find embeddings of planar graphs that satisfy certain constraints. The talk finishes with a brief discussion of extensions of SPQR-trees to planarity variants for directed graphs, such as upward- an level-planarity, where similar structures can be defined for single-source graphs.

3.2 New frontiers for graph drawing: Lossy (structural) kernelization and hybrid parameterizations

Bart M. P. Jansen (TU Eindhoven, NL)

URL https://www.win.tue.nl/ bjansen/talks/FrontiersGraphDrawing.pptx

The talk covers three frontiers for the study of graph drawing problems from the perspective of parameterized complexity. The frontiers are illustrated by examples for 2-Layer Planarization, a fruitfly for parameterized graph drawing. Additionally, a sample result on each frontier is given based on my recent work.

The first frontier consists of kernelization for structural parameters. What is the strongest parameterization that admits a polynomial kernel? This question was recently resolved for the undirected Feedback Vertex Set problem, using the notion of elimination distance.

The second frontier is lossy kernelization, which aims to reduce instances in such a way that an approximate solution to the reduced instance can be lifted to a (slightly worse) approximation to the original. A constant-factor lossy kernel of polynomial size was recently developed for Vertex Planarization.

The last frontier consists of hybrid parameterizations improving simultaneously on parameterizations by solution size and treewidth. Fixed-parameter tractable algorithms for such parameterizations were recently found for Vertex Planarization. The running time of this algorithm is not much worse than for solution-size parameterizations.

3.3 Upward Book Embeddings: Complexity and Parameterized Algorithms

Giordano Da Lozzo (University of Rome III, IT)

A *k-page upward book embedding* (kUBE) of a directed acyclic graph (a DAG, for short) is a book embedding of the DAG on k pages with the additional requirement that the vertices appear in a topological ordering along the spine of the book. The *page number* of a DAG is the minimum k for which it admits a kUBE. In 1999, Heath and Pemmaraju conjectured that the recognition problem of DAGs with page number 2, called 2UBE Testing, is NP-complete [3]. After 23 years, Bekos et al. have finally settled this conjecture in the affirmative [1]. In particular, this NP-completeness result holds even for planar st-graphs. On the algorithmic side, Binucci et al. have complemented the previous negative result by providing an $O(f(\beta) \cdot n + n^3)$-time algorithm for 2UBE Testing of planar st-graphs, where β is the branchwidth of the input graph and f is a singly-exponential function on β [2]. In this talk, we will review the previously mentioned main advancements in the complexity of 2UBE Testing. Moreover, we will survey interesting variants of the problem and highlight promising research directions in the study of parameterized algorithms for these variants.

References

1 Michael A. Bekos, Giordano Da Lozzo, Fabrizio Frati, Martin Gronemann, Tamara Mchedlidze, Chrysanthi N. Raftopoulou: Recognizing DAGs with page-number 2 is NP-complete. Theor. Comput. Sci. 946: 113689 (2023)
2 Carla Binucci, Giordano Da Lozzo, Emilio Di Giacomo, Walter Didimo, Tamara Mchedlidze, Maurizio Patrignani: Upward Book Embeddings of st-Graphs. SoCG 2019: 13:1-13:22
3 Lenwood S. Heath, Sriram V. Pemmaraju: Stack and Queue Layouts of Directed Acyclic Graphs: Part II. SIAM J. Comput. 28(5): 1588-1626 (1999)

3.4 Report from the Previous Edition: Product structure for h-framed graphs

Petr Hlinený (Masaryk University – Brno, CZ)

Joint work of Michael A. Bekos, Giordano Da Lozzo, Petr Hliněný, Michael Kaufmann
Main reference Michael A. Bekos, Giordano Da Lozzo, Petr Hlinený, Michael Kaufmann: "Graph Product Structure for h-Framed Graphs", in Proc. of the 33rd International Symposium on Algorithms and Computation, ISAAC 2022, December 19-21, 2022, Seoul, Korea, LIPIcs, Vol. 248, pp. 23:1–23:15, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022.
URL https://doi.org//10.4230/LIPIcs.ISAAC.2022.23

The focus of the working group from the 2021 seminar was on the product structure theorem of Dujmovic, Joret, Micek, Morin, Ueckerdt, and Wood (J. ACM 2020). We were following two research directions, to find nontrivial direct algorithmic applications of the product structure theory (which do not seem to exist in this strict sense), and to suitably extend the scope of the product structure beyond planarity, while still giving explicit small bounds. We succeeded in the second direction as follows.

A graph is *h-framed*, if it admits a drawing in the plane whose uncrossed edges induce a biconnected spanning plane graph with faces of size at most h. For instance, every 1-planar graph is a subgraph of a 4-framed graph, and every h-framed graph is $O(h^2)$-planar. As the main result, we have proved:

▶ **Theorem.** *Let G be a not-necessarily simple h-framed graph with $h \geq 3$. Then the simplification of G is a subgraph of the strong product $H \boxtimes P \boxtimes K_q$, where H is a planar graph with $tw(H) \leq 3$, $q = 3\lfloor \frac{h}{2} \rfloor + \lfloor \frac{h}{3} \rfloor - 1$, and P is a path.*

We have used this theorem to improve known upper bounds on the queue number and the non-repetitive chromatic number on 1-planar and optimal 2-planar graphs, and to give explicit upper bounds on the twin-width of these graphs.

3.5 Report from the Previous Edition: Upward Planarity Testing

Kirill Simonov (Hasso-Plattner-Institut, Universität Potsdam, DE)

Joint work of Kirill Simonov, Michael A. Bekos, Giordano Da Lozzo, Petr Hliněný, Michael Kaufmann
Main reference Steven Chaplick, Emilio Di Giacomo, Fabrizio Frati, Robert Ganian, Chrysanthi N. Raftopoulou, Kirill Simonov: "Parameterized Algorithms for Upward Planarity", CoRR, Vol. abs/2203.05364, 2022.
URL https://doi.org//10.48550/arXiv.2203.05364

In Upward Planarity Testing, the input is a directed acyclic graph (DAG) G and the question is whether there exists a planar drawing of G such that all edges are drawn upward, i.e., all edges monotonically increase in the vertical direction. Upward Planarity Testing has been extensively studied in the literature, and arises naturally in a number of situations where the aim is to obtain easy-to-parse planar representations of DAGs. The problem has been shown to be NP-complete already 25 years ago [5], and there are a few special cases known when the problem can be solved in polynomial time. Among others, when G is restricted to class of single-source DAGs [1], or the class of orientations of series-parallel graphs [4]. There have also been a couple of fixed-parameter tractability results for Upward Planarity Testing.

It was the goal of our working group at the previous edition of the seminar to push the boundaries Upward Planarity Testing from the parameterized complexity perspective, generalizing and unifying the previous results. Our first target was to generalize the polynomial-time single-source algorithm [1], which resulted in a single-exponential FPT algorithm when parameterized by the number of sources [2]. We also considered the problem parameterized by prominent structural parameters. We showed that Upward Planarity Testing is XP parameterized by treewidth, and FPT parameterized by treedepth of the underlying undirected graph [2]. All three algorithms are obtained using a novel framework for the problem that combines SPQR tree-decompositions with parameterized techniques. Finally, we used our improved analysis on the series-parallel case: We showed that Upward Planarity Testing can be solved in time $O(n^2)$ on n-vertex series-paralled graphs [3], while the previous-best algorithm runs in time $O(n^4)$ [4].

References

1 Paola Bertolazzi, Giuseppe Di Battista, Carlo Mannino, Roberto Tamassia: *Optimal Upward Planarity Testing of Single-Source Digraphs*. SIAM J. Comput. 27(1): 132-169 (1998)
2 Steven Chaplick, Emilio Di Giacomo, Fabrizio Frati, Robert Ganian, Chrysanthi N. Raftopoulou, Kirill Simonov: *Parameterized Algorithms for Upward Planarity*. SoCG 2022: 26:1–26:16

3 Steven Chaplick, Emilio Di Giacomo, Fabrizio Frati, Robert Ganian, Chrysanthi N. Raftopoulou, Kirill Simonov: *Testing Upward Planarity of Partial 2-Trees*. GD 2022: 175-187
4 Walter Didimo, Francesco Giordano, Giuseppe Liotta: *Upward Spirality and Upward Planarity Testing*. SIAM J. Discret. Math. 23(4): 1842-1899 (2009)
5 Ashim Garg, Roberto Tamassia: *On the computational complexity of upward and rectilinear planarity testing*. SIAM J. Comput. 31(2):601–625 (2001)

3.6 Report from the Previous Edition: A Parameterized Approach to Orthogonal Compaction

Siddharth Gupta (University of Warwick – Coventry, GB)

Joint work of Siddharth Gupta, Philipp Kindermann, Walter Didimo, Giuseppe Liotta, Alexander Wolff, Meirav Zehavi
Main reference Walter Didimo, Siddharth Gupta, Philipp Kindermann, Giuseppe Liotta, Alexander Wolff, Meirav Zehavi: "Parameterized Approaches to Orthogonal Compaction", CoRR, Vol. abs/2210.05019, 2022.
URL https://doi.org//10.48550/arXiv.2210.05019

Orthogonal graph drawings are used in applications such as UML diagrams, VLSI layout, cable plans, and metro maps. We focus on drawing planar graphs and assume that we are given an *orthogonal representation* that describes the desired shape, but not the exact coordinates of a drawing. Our aim is to compute an orthogonal drawing on the grid that has minimum area among all grid drawings that adhere to the given *orthogonal representation*.

This problem is called orthogonal compaction (OC) and is known to be NP-hard, even for orthogonal representations of cycles [1]. We investigate the complexity of OC with respect to several parameters. Among others, we show that OC is fixed-parameter tractable with respect to the most natural of these parameters, namely, the number of *kitty corners* of the orthogonal representation: the presence of pairs of kitty corners in an orthogonal representation makes the OC problem hard. Informally speaking, a pair of kitty corners is a pair of reflex corners of a face that point at each other. Accordingly, the number of kitty corners is the number of corners that are involved in some pair of kitty corners.

References

1 W. S. Evans, K. Fleszar, P. Kindermann, N. Saeedi, C.-S. Shin, and A. Wolff. Mini- mum rectilinear polygons for given angle sequences. Comput. Geom., 100(101820):1– 39, 2022. doi:10.1016/j.comgeo.2021.101820.

4 Open problems

4.1 The Maximum Bimodal Subgraph Problem

Walter Didimo (University of Perugia, IT)

Let G be a plane digraph, i.e., a directed planar graph with a fixed planar embedding. A vertex v of G is bimodal if all its incoming edges (and hence all its outgoing edges) are consecutive in the cyclic order around v. The graph G is bimodal if all its vertices are bimodal. When G is not bimodal, the *maximum bimodal subgraph* problem (MBS) asks to extract from G an embedding preserving subgraph with the maximum number of edges. This problem is known to be NP-hard by Binucci et al. [1], who also describe an efficient heuristic and a branch-and-bound expnential-time algorithm.

The goal is to initiate the study of the MBS problem from the parameterized complexity perspective. In particular, investigate the existance of a polynomial kernel of the MBS parameterized by the number of non-bimodal vertices. Also, is the MBS problem in FPT for some of its structural parameterizatons?

References

1 Carla Binucci, Walter Didimo, and Francesco Giordano. Maximum upward planar subgraphs of embedded planar digraphs. *Comput. Geom.*, 41(3):230–246, 2008.

2 Frederic Dorn, Eelko Penninkx, Hans L. Bodlaender, and Fedor V. Fomin. Efficient exact algorithms on planar graphs: Exploiting sphere cut decompositions. *Algorithmica*, 58(3):790–810, 2010.

3 Brenda S. Baker. Approximation algorithms for NP-complete problems on planar graphs. *J. Assoc. Comput. Mach.*, 41(1):153–180, 1994.

4.2 Planar $\mathcal{F}$-Augmentation of Clustered Graphs

Giordano Da Lozzo (University of Rome III, IT) and Siddharth Gupta (University of Warwick – Coventry, GB)

A *flat clustered graph* (for short *flat c-graph*) is a pair $(G, \mathcal{V})$ where G is a graph and $\mathcal{V}$ is a partition of the vertices of G. The parts of $\mathcal{V}$ are called *clusters*. A flat c-graph is *independent* if each of its clusters induces an independent set. Given a family $\mathcal{F}$ of connected graphs and an independent flat c-graph $\mathcal{C} = (G, \mathcal{V} = \{V_1, \dots, V_k\})$, the Planar $\mathcal{F}$-Augmentation problem for $\mathcal{C}$ asks to test for the existence of sets $Z_1, \dots, Z_k$ of non-edges of G such that $G \cup \bigcup_{i=1}^{k} Z_i$ is planar and $G(V_i, Z_i) \in \mathcal{F}$ for $i = 1, \dots, k$. If $\mathcal{F}$ is the family of all paths, then the problem coincides with the Clustered Planarity with Linear Saturators problem, which is known to be NP-complete [1]. On the other hand, if $\mathcal{F}$ is the family of all trees, then the problem coincides with the Clustered Planarity problem, which has been recently shown to be solvable in polynomial time [2].

We highlight some interesting research directions in the study of the Planar $\mathcal{F}$-augmentation problem. First, concerning the setting in which $\mathcal{F}$ is the set of all paths, the mentioned NP-completeness stimulates the search for exact parameterized and sub-exponential algorithms for the problem, both in the fixed and in the variable embedding setting. Moreover, albeit the problem is known to be NP-complete if all clusters, except one, are *trivial*, i.e., they each consist of a single vertex, it is still unknown whether the problem is para-NP-hard in the total number of clusters. Finally, as paths and trees may have linear radius, we regard as an appealing question, at the other side of the radius spectrum, to settle the complexity of the problem when $\mathcal{F}$ is the family of all stars. In this latter setting, it is possible to show that the problem reduces to a constrained planarity problem on the graph obtained by identifying all the vertices belonging to the same cluster, whose study may be of independent interest.

References

1 Patrizio Angelini, Giordano Da Lozzo, Giuseppe Di Battista, Fabrizio Frati, Maurizio Patrignani, Ignaz Rutter: *Intersection-Link Representations of Graphs*. J. Graph Algorithms Appl. 21(4): 731-755 (2017)

2 Radoslav Fulek, Csaba D. Tóth: *Atomic Embeddability, Clustered Planarity, and Thickenability*. J. ACM 69(2): 13:1-13:34 (2022)

4.3 The complexity of recognizing Witness Graphs

Eduard Eiben (Royal Holloway, University of London, GB)

Main reference Boris Aronov, Muriel Dulieu, Ferran Hurtado: "Witness (Delaunay) graphs", Comput. Geom., Vol. 44(6-7), pp. 329–344, 2011.
URL https://doi.org//10.1016/j.comgeo.2011.01.001

Witness graphs are a generalization of proximity graphs introduced by Aronov, Dulieu, and Hurtado [1]. A witness graph $G = (V, E)$ is defined by a quadruple $(P, S, W, \pm)$: P is a set of points in plane and $V = P$; S provides the geometric shapes (disks, squares, rectangles); W is set of "witness" points/witnesses in plane; and $uv \in E$ if and only if

- if $+$: some $s \in S$ covers u, v, w for some $w \in W$
- if $-$: some $s \in S$ covers u, v, but no $w \in W$

Examples of witness graphs include:

- Any Delaunay Triangulation can be defined as $(P, S, P, -)$, with S being the set of all disks.
- $(P, S, \emptyset, -)$ defines a complete graph on P as the vertex set.
- Witness Gabriel Graphs [2] are defined by $(P, S, W, -)$, where S contains disks for pairs $u, v \in P$, such that uv is diameter.

Open Questions

What is the complexity of the following problems:

1. Given a graph G and positive integer k, can we represent it as Witness Delaunay Graph/W. Gabriel Graph/Witness Square Graph with at most k witnesses (i.e., $|W| \leq k$)?
2. Given graph G, the set of points P, and S, can we find W such that G is witness graph represented by $(P, S, W, -)$ or $(P, S, W, +)$? A polynomial time algorithm is known for Witness Gabriel Graph. However, the complexity of finding a minimum size W is open even for Witness Gabriel Graph.

References

1 B. Aronov, M. Dulieu, and F. Hurtado. Witness (delaunay) graphs. *Comput. Geom.*, 44(6-7):329–344, 2011.

2 B. Aronov, M. Dulieu, and F. Hurtado. Witness gabriel graphs. *Comput. Geom.*, 46(7):894–908, 2013.

4.4 Depth-first search trees of minimum height

Petr A. Golovach (University of Bergen, NO)

Spanning trees, in particular, depth-first spanning trees (DFS-trees) are popular tools in graph drawing. For these purposes, it may be convenient to consider DFS-trees with additional properties like minimum/maximum number of leaves or minimum height. The additional constraints usually make finding such trees NP-hard. The investigation of the parameterized complexity of problems of this type was initiated by Fellows, Friesen, and Langston [1]. We are interested in the parameterized complexity of Min Height DFS-Tree problem whose task is, given a graph G and an integer h, to decide whether G has a DFS-tree of height at most h. This problem is known to be NP-complete [1]. It can be observed that if G has a DFS-tree of height at most h then the treedepth of G is upper bounded by h. This implies that Min Height DFS-Tree is FPT when parameterized by h. Is it possible to give an efficient, say, single-exponential direct algorithm for the problem? Another direction of research is to investigate the (parameterized) approximability of the minimum height of a DFS-tree.

References

1 M. R. Fellows, D. K. Friesen, and M. A. Langston, *On Finding Optimal and Near-Optimal Lineal Spanning Trees*, Algorithmica, 3 (1988), pp. 549–560.

4.5 Sequential Crossing Number

Thekla Hamm (Utrecht University, NL)

One of the most classically important features of graph drawings is a small number of crossings and the Crossing Number problem, which asks for the minimum number of crossings a graph can be drawn with , is arguably the most famous graph drawing problem. A driving motivation of the field of graph drawing is the visualisation of graphs – a task which is ubiquitous in areas such as the analysis of social networks, bioinformatics and linguistics or generally when information can be captured by graphs. Often such information can evolve over time and hence the treatment of temporal aspects is an important challenge in graph drawing. In particular, is reasonable to require some "stability" of drawing of a graph that changes over time in the sense that parts of the graph occur in consecutive time steps are not redrawn in a different way.

Motivated by this, we define the Sequential Crossing Number problem as generalisation of Crossing Number in which, given a sequence of graphs, we want to draw each of them in a way that drawings of intersections of consecutive graphs agree, and the total number of crossings (i.e. the sum of the number of crossings in each drawing) is minimised.

For $i \in [t-1]$ we say that G_i and G_{i+1} are *consecutive*. Let us also point out two alternative objectives that could be motivated in the same spirit: Instead of requiring $\sum_{i\in[t]} \mathtt{crn}(\mathcal{G}_i) \leq k$, it could also be interesting to require $\max_{i\in[t]} \mathtt{crn}(\mathcal{G}_i) \leq k$, i.e. minimise the maximum number of crossings in any graph, or $\sum_{i\in[t]} \mathtt{crn}(\mathcal{G}_i) - \mathtt{crn}(\mathcal{G}_i|_{G_i\cap G_{i-1}}) \leq k$, i.e. minimise the total number of crossings but counting crossings the drawing of the intersection of consecutive graphs only once.

Of course, classic Crossing Number is captured by $t = 1$, and this already implies NP-hardness of the problem. This is why our open problems will target parameterised algorithms for Sequential Crossing Number. Unless specified otherwise, we consider the parameterisation by k.

Sequential Crossing Number is closely related to the Simultaneous Crossing Number problem [1] which was proposed as a natural generalisation of the well-known Simultaneous Embedding with Fixed Edges problem (t-SEFE for short and most prominently considered for $t = 2$) [2], but has not received a lot of attention since. The conceptual difference between Sequential Crossing Number and Simultaneous Crossing Number is that the former ensures stability of the drawings of subgraphs as long as they appear, whereas the latter additionally ensures stability of drawings of subgraphs that may disappear and reappear over time. t-SEFE is equivalent to Simultaneous Crossing Number for $k = 0$.

Both problems are, by definition equivalent as soon as there is some graph H such that $\forall i \in [t-1]\ G_i \cap G_{i+1} = H$. In particular, this means that the known NP-hardness for t-SEFE on such instances, even if $t = 3$ and H is a star implies that we cannot hope for a parameterised (even XP) algorithm for Sequential Crossing Number in $k + t$, even on instances with a common pairwise intersection of consecutive graphs which is a star.

We pose the general open problem:

▶ **P 1.** *Can we identify conditions on the pairwise intersections of consecutive graphs under which there is an* FPT *or* XP*-algorithm for Sequential Crossing Number in k or $k + t$?*

We propose two concrete directions for such conditions: In the special case that $k = 0$ and each intersection of consecutive graphs is 3-connected, the Sequential Crossing Number is equivalent to testing the planarity of each graph, and hence polynomial-time solvable. Moreover, branching and rigidising the drawings of intersections of consecutive graphs which contain crossings and branching on the graphs whose drawings contain crossings involving edges outside of any consecutive intersection in a hypothetical solution reduces Sequential Crossing Number on instances in which the pairwise intersections of consecutive graphs are 3-connected to Partially Predrawn Crossing Number (the problem of deciding the minimum number of crossings required to extend a drawing of a subgraph to the entire graph). Partially Predrawn Crossing Number is known to be in FPT [3]. From this it can be argued that Sequential Crossing Number on instances in which the pairwise intersections of consecutive graphs are 3-connected is in XP.

▶ **P 2.** *Is Sequential Crossing Number on instances in which the pairwise intersections of consecutive graphs are 3-connected in* FPT*?*

A similar idea would be to rigidise branched drawings for small intersections of consecutive graphs which could yield an XP-algorithm for Sequential Crossing Number parameterised by $k + t + \max_{i \in [t-1]} |E(G_i \cap G_{i+1})|$.

▶ **P 3.** *Is Sequential Crossing Number in* FPT *parameterised by $k+t+\max_{i \in [t-1]} |E(G_i \cap G_{i+1})|$?*

In a different direction polynomial-time algorithm known for t-SEFE on instances with a common pairwise intersection of all input graphs requires that the pairwise intersection of graphs be 2-connected. It would be interesting to extend this result:

▶ **P 4.** *Is Sequential Crossing Number on instances with a common pairwise 2-connected intersection of consecutive graphs in* FPT*? Is this even in* XP*? What about the parameterisation by $(k + t)$?*

References

1 M. Chimani, M. Junger and M. Schulz, "Crossing Minimization meets Simultaneous Drawing", 2008 IEEE Pacific Visualization Symposium, 2008, pp. 33-40, doi: 10.1109/PACIFICVIS.2008.4475456.

2 I. Rutter, "Simultaneous Embedding", Beyond Planar Graphs, 2020, doi: 10.1007/978-981-15-6533-5_13

3 T. Hamm and P. Hliněný, "Parameterised Partially-Predrawn Crossing Number", 38th International Symposium on Computational Geometry (SoCG 2022), 2022, doi: 10.4230/LIPIcs.SoCG.2022.46

4.6 On Left-Aligned BFS Trees

Petr Hlinený (Masaryk University – Brno, CZ)

Main reference Petr Hlinený, Jan Jedelský: "Twin-width of Planar Graphs is at most 8", CoRR, Vol. abs/2210.08620, 2022.
URL https://doi.org//10.48550/arXiv.2210.08620

The concept of a *left-aligned BFS tree* of a planar graph, defined below, was formulated by P. Hliněný while working on an upper bound on the twin-width of planar graphs. It is an important ingredient in the currently best such bound of 8 by Hliněný and Jedelský (briefly noting, without left-aligned BFS trees the same proof would yield only an upper bound of 11). A left-aligned BFS tree can be computed in linear time for a given planar embedding. The featured question (or open problem) is whether the concept of a left-aligned BFS tree can find other applications in algorithmic problems related to planarity, and in particular in the graph drawing area.

▶ **Definition.** *Consider a plane graph G, and a BFS tree T spanning G and rooted in a vertex r of the outer face of G, and picture (for clarity) the embedding G such that r is the vertex of G most at the top. An informal intention is to have a BFS tree T such that there is no other BFS tree of G which is "to the left" of T in at least some place of the geometric picture of G and T. Formally, for two adjacent vertices $u, v \in V(G)$, $\{u, v\} \in E(G)$, we say that u is to the left of v (wrt. T) if neither of u, v lies on the vertical path from r to the other, and the following holds; if r' is the least common ancestor of u and v in T and $P_{r',u}$ (resp., $P_{r',v}$) denote the vertical path from r' to u (resp., v), then the cycle $(P_{r',u} \cup P_{r',v}) + uv$ has the triple (r', u, v) in this counter-clockwise cyclic order. A BFS tree T of G with the BFS layering $\mathcal{L} = (L_0, L_1, \ldots)$ is called* left-aligned *if there is no edge $f = uv$ of G such that, for some index i, $u \in L_{i-1}$ and $v \in L_i$, and u is to the left of v.*

We remark that this concept might seem closely related to the previously used leftist canonical ordering of planar graphs by Badent, Baur, Brandes, and Cornelsen, but it is not that close since the latter does have have the BFS property. Yet, possible applications of both notions may lie in similar areas.

4.7 The *st*-Planar Graph Augmentation Problem

Fabrizio Montecchiani (University of Perugia, IT)

A directed graph G is an st-planar graph if it admits a planar embedding such that: (i) it contains no directed cycle; (ii) it contains a single source vertex s and a single sink vertex t; and (iii) s and t both belong to the outer face of the planar embedding. A popular result in graph drawing states that a directed acyclic graph (DAG) G has an upward planar drawing if and only if G is a subgraph of an st-planar graph. Since testing upward planarity is NP-complete, it follows that testing if a graph is a subgraph of an st-planar graph is also hard. On the other hand, checking if a DAG is st-planar can be easily done in polynomial time.

Based on this motivation we propose to study the st-Planar Graph Augmentation Problem (stPA), which is defined as follows. Let $G = (V, E)$ be a DAG, and let k be a positive integer. Is it possible to *add* at most k edges to G such that the resulting graph is an st-planar graph up to the addition of a super source s and of a super sink t in the outer face of the graph? As a natural open problem, we ask whether stPA admits a fixed-parameter tractable algorithm when parameterized in its natural parameter k.

4.8 The parameterized complexity of Queue Number using structural parameters

Sebastian Ordyniak (University of Leeds, GB)

The Queue Number of an undirected graph G is the minimum number of pages h such that G has s h-queue layout (see also `https://en.wikipedia.org/wiki/Queue_number`).

An h-queue layout of $G = (V, E)$ is an ordering $<_q$ of the vertices together with a partition $\mathcal{P}$ of the edge set of G into h sets such that for every $P \in \mathcal{P}$, if the vertices of G are drawn on the line according to the ordering $<_q$, then no two edges in P form a rainbow, i.e., there are no two edges$e = (u, v)$ and $e = (x, y)$ in P such that $u < x < y < v$. The following is known about Queue Number:

- NP-complete even for one page.
- polynomial-time for fixed ordering (due to k-rainbow obstructions)
- For one page: known to be fpt by treedepth [1]
- FPT by vertex cover for arbitrary number of pages

Open Problems:

- For one page: FPT by treewidth or feedback edge number?
- For arbitrary number of pages: FPT by feedback edge number, treedepth, treewidth?

References

1 Sujoy Bhore, Robert Ganian, Fabrizio Montecchiani and Martin Nöllenburg, *Parameterized Algorithms for Queue Layouts*. J. Graph Algorithms Appl., vol. 26, number 3, pp. 335–352, 2022

4.9 Upward Planarity Testing Parameterized by Treewidth

Kirill Simonov (Hasso-Plattner-Institut, Universität Potsdam, DE)

In Upward Planarity Testing, the input is a directed acyclic graph (DAG) G and the question is whether there exists a planar drawing of G such that all edges are drawn upward, i.e., all edges monotonically increase in the vertical direction. Upward Planarity Testing has been extensively studied, and was shown to be NP-complete already 25 years ago [1], while many polynomial-time algorithms in special cases were also known.

As an outcome of the previous edition of the seminar, it has been shown that Upward Planarity Testing admits an algorithm with running time $n^{O(\text{tw})}$ [2], where n is the number of vertices in the graph and tw is its treewidth. In parameterized complexity terms, Upward Planarity Testing is thus in XP when parameterized by the treewidth of the underlying undirected graph.

A natural question is then whether the algorithmic result above can be improved. Specifically, is Upward Planarity Testing W[1]-hard or FPT parameterized by treewidth? If the former holds, is the running time of $n^{O(\text{tw})}$ tight. For example, can it be shown that there is no $n^{o(\text{tw})}$-time algorithm under the ETH?

References

1 Ashim Garg, Roberto Tamassia: *On the computational complexity of upward and rectilinear planarity testing*. SIAM J. Comput. 31(2):601–625 (2001)
2 Steven Chaplick, Emilio Di Giacomo, Fabrizio Frati, Robert Ganian, Chrysanthi N. Raftopoulou, Kirill Simonov: *Parameterized Algorithms for Upward Planarity*. SoCG 2022: 26:1–26:16

4.10 Graph embeddings: Drawing vs. learning

Manuel Sorge (TU Wien, AT)

Graph drawing is about finding embeddings of graphs in low-dimensional space. I proposed to look at the embedding problem from a new angle, namely, from a machine-learning point of view. In machine learning, embeddings are also mappings into a low-dimensional space, that represents a condensed form of information, from a high-dimensional space, such as the space of human-written documents or the space of graphs from an application area. The goals of the classical embedding in graph drawing and embedding in machine learning are related, but different: In graph drawing, we usually aim for readability by, e.g., minimizing the number of crossings. In machine learning, we aim to embed similar entities (such as Dagstuhl open problem descriptions) close to each other in the embedding space. A famous embedding algorithm in the machine-learning realm is Node2vec [1], which aims to map the vertex set of a given graph to a low-dimensional embedding space such that similar vertices, as measured by a certain neighborhood relation based on co-occurrence in random walks, are close to each other in the embedding space.

It seems interesting to:

1. Explore machine-learning embedding problems from a visualization point of view, i.e., is a graph easier to understand if we embed similar vertices close to each other in 2d-space, like in Node2vec?
2. Consider discrete variants of embedding problems: The dimensions in the embedding space can be thought of as underlying features of the data. For interpretability, it would be desirable to have few easy-to-understand features. I.e., it would be interesting to map to low-dimensional spaces where each dimension has only small domain, that is, a small number of admissible values.
3. Develop efficient algorithms that exploit properties such as the low dimensionality of the embedding space, small domains, and input graph structure.

For a concrete starting point, we may look at a simplified version of Node2vec: Here, we are given a graph $G = (V, E)$ and we aim to find a mapping $f \colon V \to \mathbb{R}^d$ for some small d such that $\sum_{u \in V} \log \Pr(N(u) \mid f(u))$ is maximized. Intuitively and simplified, we want to maximize the number of vertices with the same neighborhood that are mapped to the same point in the embedding space. After several approximations, Grover and Leskovec [1] arrive at the objective to maximize

$$\sum_{u \in V} \left[-\log(Z_u) + \sum_{v \in N_S(u)} f(u)f(v) \right],$$

where $Z_u = \sum_{v \in V} \exp(f(u)f(v))$. Herein, we can think of the logarithmic term as a baseline and each term in the sum measures the deviation from the baseline. Simplifying further, we may observe that in the vertices in the baseline term are weighed lower (due to taking the logarithm) than the vertices in the other term. If we weigh both at exactly half instead, we obtain the following objective:

$$\sum_{u \in V} \left[\sum_{v \in N(u)} (f(u)f(v)) - \sum_{v \in V \setminus N(u)} f(u)f(v) \right].$$

For instance, if we simply want to find an embedding into binary, one-dimensional space $\{0, 1\}$, we thus want to find an induced subgraph (corresponding to the vertices assigned 1) that maximizes the number of edges contained in the subgraph minus the number of nonedges in the subgraph. A natural first step is to find some baseline complexity results or promising algorithms by looking at such special cases and then generalize to larger domains and dimensions.

References

1 Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In Balaji Krishnapuram, Mohak Shah, Alexander J. Smola, Charu C. Aggarwal, Dou Shen, and Rajeev Rastogi, editors, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*, pages 855–864. ACM, 2016. ISBN 978-1-4503-4232-2. doi: 10.1145/2939672.2939754. URL `https://doi.org/10.1145/2939672.2939754`.

4.11 Drawing Hypergraphs "Nicely"

Yushi Uno (Osaka Metropolitan University, JP)

Hypergraphs [1] are not only theoretically interesting as a generalized notion of graphs, but also has a wide range of applications in many fields, such as database theory, biology, and so on [3]. Therefore, drawing and visualizing hypergraphs is useful for understanding practical cases in applied fields.

Examples of such visualized hypergraphs have been appeared in typical textbooks or papers [1, 2], however, there are many different ways to draw them, and a unified approach has not yet been fully established. On the other hand, it is often experienced that the level of understanding a hypergraph varies greatly depending on how it is drawn, and there is a need for "nice" drawing of hypergraphs.

The criteria for goodness or understandability of a hypergraph drawing could simply be the number of crossings of hyperedges or the number of curved lines, for example, but the definition of the crossings itself is not necessarily trivial that might be depending on the drawing model. Thus, for example, the following kind of problem can be considered:

<u>Minimum Number of Curved Lines</u>
Input: a k-uniform hypergraph ($k \geq 3$),
Output: under "straight and curved line model", what is the minimum number of curved lines, and what is its node layout?

We would then like to discuss drawing algorithms that optimize such criteria and analyze these algorithms in terms of their parameterized computational complexity. Possible parameters include the number of hyperedges, the number of crossings, the maximum degree, uniformity, and so on.

This topic is closely related to the notion of set visualization, e.g., `https://www.dagstuhl.de/en/seminars/seminar-calendar/seminar-details/22462`, so we have to scrutinize the results in that field.

References

1 C. Berge. *Graphs and Hypergraphs.* North-Holland, 1973.
2 V. I. Voloshin. *Introduction to Graph and Hypergraph Theory.* Nova Science, 2009.
3 R. Fagin. Degrees of acyclicity for hypergraphs and relational database schemes. *J. ACM*, 30(3), 514–550, 1983.

5 Working groups

5.1 A Subexponential FPT Algorithm for Deleting Few Edges from Ordered Graphs to Make Them p-Plane

Akanksha Agrawal (Indian Institute of Techology Madras, IN), Sergio Cabello (University of Ljubljana, SI), Michael Kaufmann (Universität Tübingen, DE), Saket Saurabh (The Institute of Mathematical Sciences – Chennai, IN), Roohani Sharma (MPI für Informatik – Saarbrücken, DE), Yushi Uno (Osaka Metropolitan University, JP), and Alexander Wolff (Universität Würzburg, DE)

Introduction

Since (many) crossings usually make it hard to understand the drawing of a graph, much effort in the area of Graph Drawing has been directed towards reducing the number of crossings in drawings of graphs. In terms of parametrized complexity, several facets of this problem have been considered. For example, there are FPT algorithms that, given a graph G and an integer k, decide whether G can be drawn with at most k crossings [8, 12].

Crossing minimization has also been considered in the setting where each vertex of the given graph must lie on one of two horizontal lines. This restricted version of crossing minimization is an important subproblem in drawing layered graphs according to the so-called *Sugiyama framework* [17]. There are two variants of the problem; either the vertices on both lines may be freely permuted or the order of the vertices on one line is given. These variants are called two-layer and one-layer crossing minimization, respectively. For both, FPT algorithms exist [5, 13].

Surprisingly, crossing minimization remains NP-hard even when restricted to graphs that have a planar subgraph with just one edge less [2]. Another way to deal with crossings is to remove a small number of vertices or edges such that the remaining graph can be drawn without crossings. In fact, it is known that vertex deletion to planarity is FPT with respect to the number of deleted vertices [9, 11, 15]. However, the running times of these algorithms depends at least exponentially on the number of deleted vertices. On the kernelization front, there exists an α-approximate kernel for vertex deletion to planarity [10], whereas vertex deletion to outerplanarity is known to admit an (exact) polynomial kernel [4].

In the following, we consider a restricted setting that allows us to obtain substantially better running times although the graphs to which we delete are more complex than (outer-)planar graphs.

An *ordered graph* (G, σ) consists of a graph G and an ordering $\sigma\colon V(G) \to \{1, \dots, |V(G)|\}$ of the vertices of G. For two edges $e = (u, v)$ and $e' = (u', v')$ of an ordered graph (G, σ), we say that e and e' *cross with respect to* σ if $\sigma(u) < \sigma(u') < \sigma(v) < \sigma(v')$ or if $\sigma(u') < \sigma(u) < \sigma(v') < \sigma(v)$. Let $\texttt{cross}_{(G,\sigma)}(e)$ denote the set of edges of G that cross e with respect to σ. We drop the subscript (G, σ) when it is clear from the context. The ordered graph models the scenario where the vertices of G are placed along a horizontal line in the order given by σ and all the edges are drawn above the line using curves that cross as few times as possible. Whenever e and e' cross with respect to σ, their curves must intersect. Whenever e and e' do *not* cross with respect to σ, their curves can be drawn without intersections; for example, we may use halfcircles. In this setting, we get a drawing such that two edges of G cross precisely if and only if they cross with respect to σ.

We consider the edges of G as directed from the vertex with smaller index to that of larger index, that is, edge uv is directed (u, v) if $\sigma(u) < \sigma(v)$. Given a positive integer p, we say that an ordered graph (G, σ) is p-plane if every edge in G is crossed by at most p other edges. We study the parameterized complexity of the following problem.

Ordered Graph Edge Deletion to p-Plane	**Parameter:** k
Input: An ordered graph (G, σ), and a positive integer k	
Question: Does there exist a set $S \subseteq E(G)$ of at most k edges such that $(G - S, \sigma)$ is p-plane.	

An instance of this problem is denoted $((G, \sigma), k)$ or simply (G, σ, k).

Given an ordered graph (G, σ), define the *conflict graph* $H_{(G,\sigma)}$ as the graph that has a vertex for each edge of G and an edge for each pair of crossing edges of G. Using this notion, we can express Ordered Graph Edge Deletion to p-Plane equivalently as the problem of deleting from $H_{(G,\sigma)}$ a set of at most k vertices such that the remaining graph has maximum degree at most p. For general graphs, this problem is called Vertex Deletion to Degree-p [16]; it admits a quadratic kernel [7, 18].

A Fast FPT Algorithm

The main result of this report is as follows.

▶ **Theorem 1.** *Ordered Graph Edge Deletion to p-Plane admits an algorithm with running time* $2^{\mathcal{O}(p\sqrt{k}\log(p+k))} \cdot n^{\mathcal{O}(1)}$.

In other words, we obtain a subexponential fixed-parameter tractable algorithm parameterized for Ordered Graph Edge Deletion to p-Plane parameterized by k; note that we consider p to be a constant here (although we made explicit how the running time depends on p). Our algorithm to prove Theorem 1 has two steps. First it branches on edges that are crossed by at least $p + \sqrt{k}$ other edges. When such edges do not exist, we show that the conflict graph $H_{(G,\sigma)}$ has treewidth $O(p + \sqrt{k})$, which allows us to use a known (folklore) algorithm [14] for Vertex Deletion to Degree-p whose dependency is singly exponential in the treewidth of $H_{(G,\sigma)}$.

Branching

We show that we can use branching to reduce any instance to a collection of instances where each edge e of the graph satisfies $|\texttt{cross}(e)| < p + \sqrt{k}$.

Let e be an edge of G with $|\texttt{cross}(e)| \geq p + \lceil\sqrt{k}\rceil$. If $|\texttt{cross}(e)| > p + k$, then e must be deleted, as we cannot afford to keep e and delete enough edges from $\texttt{cross}(e)$. If $|\texttt{cross}(e)| \leq p + k$, then either e must be deleted or at least $|\texttt{cross}(e)| - p$ many edges from $\texttt{cross}(e)$ must be deleted, so that at most p edges of $\texttt{cross}(e)$ stay. This results in the following branching rule, where we return an OR over the answers of the following instances:

1. Recursively solve the instance $(G - e, \sigma, k - 1)$. This branch is called the *light* branch.
2. If $|\texttt{cross}(e)| > p + k$, we do not consider other branches. Otherwise, for each subset X of $\texttt{cross}(e)$ with $|\texttt{cross}(e)| - p$ many edges, recursively solve the instance $(G - X, \sigma, k - |X|)$. Each of these branches are called *heavy* branches.

We are going to show that the recursion tree has $2^{O(p\sqrt{k}\log(p+k))}$ branches. Note that we have

$$\binom{|\texttt{cross}(e)|}{|\texttt{cross}(e)|-p} \;\leq\; \binom{|\texttt{cross}(e)|}{p} \;\leq\; \binom{p+k}{p} \;\leq\; (p+k)^p$$

possible heavy branches at each node. Let $\kappa = (p+k)^p$.

To prove the desired upper bound we interpret the branching tree as follows. First note that, in each node, we have at most κ heavy branches. We associate a distinct word over the alphabet $\Sigma = \{0, 1, \ldots, \kappa\}$ to each leaf (root to leaf path) of the recurrence tree. For each node of the recurrence tree, associate a character from Σ in each of its children such that the child node corresponding to the light branch gets appended the character 0 and the other nodes (corresponding to the heavy branches) get a distinct character from Σ. Now a word over the alphabet Σ for a leaf ℓ of the recurrence tree is obtained by taking the sequence of character on the nodes of the root to leaf ℓ, in order. In order to bound the number of leaves (and hence the total number of nodes) of the recurrence tree, it is enough to bound the number of such words. The character 0 is called a *light* label and all other characters are called *heavy* labels. Recall that a light label corresponds to the branch where k drops by 1, while the heavy labels correspond to the branches where k drops by $|\texttt{cross}(e)| - p \geq \sqrt{k}$. This implies that each word (that is associated with the leaf of the recurrence tree) has at most $\sqrt{k}$ heavy labels. In order to bound the number of such words, we first guess the places in the word that are occupied by heavy labels and then we guess the (heavy) labels itself at these selected places. All other positions have the light label on them and there is no choice left. Hence, the number of such words is upper bounded by

$$\sum_{i=0}^{\sqrt{k}} \binom{k}{i}\kappa^i \;\leq\; \sqrt{k}\binom{k}{\sqrt{k}}\kappa^{\sqrt{k}} \;\leq\; \sqrt{k}\binom{k}{\sqrt{k}}((p+k)^p)^{\sqrt{k}} \;=\; 2^{\mathcal{O}(p\sqrt{k}\log(p+k))}.$$

This shows that the number of such words are upper bounded by $2^{\mathcal{O}(p\sqrt{k}\log(p+k))}$, and hence the number of leaves (and nodes) of the recurrence tree are upper bounded by $2^{\mathcal{O}(p\sqrt{k}\log(p+k))}$.

Balanced Separators

Let (G, σ) be an ordered graph. For any edge $e = (u, v)$ of G, let $\texttt{span}_{(G,\sigma)}(e)$ be the set of all edges $(u', v') \neq e$ of G such that $\sigma(u) \leq \sigma(u') \leq \sigma(v') \leq \sigma(v)$. For any vertex w of G, let $\texttt{left}_{(G,\sigma)}(w)$ be the set of all edges (u, v) of G such that $\sigma(u) < \sigma(v) \leq \sigma(w)$. Whenever it is clear from the context, we will drop the subscript (G, σ). We say that an edge e of G is *maximal* if G contains no edge e' such that $e \in \texttt{span}(e')$,

▶ **Lemma 2** (Balanced Separator). *If (G, σ) is an ordered q-plane graph, then G contains a set X of at most $2(q + 1)$ edges such that $E(G) \setminus X = E_1 \uplus E_2$, $|E_1| \leq 2m/3$, $|E_2| \leq 2m/3$, and no edge $e_1 \in E_1$ crosses an edge $e_2 \in E_2$ with respect to σ.*

Proof. We consider three cases depending on the spans of the edges of G.

Case 1: There exists an edge $e = (u, v) \in E(G)$ such that $m/3 \leq |\texttt{span}(e)| \leq 2m/3$.

In this case, let $X = \texttt{cross}(e) \cup \{e\}$, let $E_1 = \texttt{span}(e)$, and let $E_2 = E(G - X) \setminus E_1$. Note that $|X| \leq q + 1$, $|E_1| \leq 2m/3$, and $|E_2| \leq 2m/3$. Now let $e_1 = (u_1, v_1) \in E_1$ and

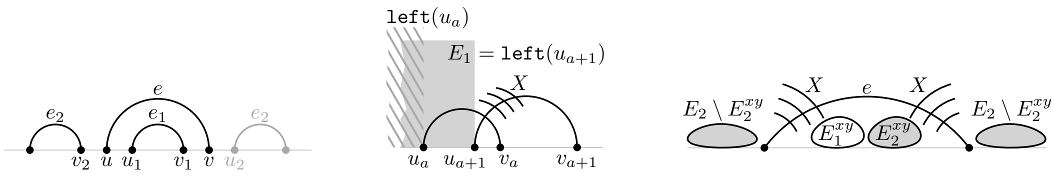

Figure 1 Case distinction for the proof of theorem 2.

$e_2 = (u_2, v_2) \in E_2$. Since $e_1 \in E_1$, we have $\sigma(u) \leq \sigma(u_1) < \sigma(u_2) \leq \sigma(v)$; see Case 1, fig. 1. Since $e_2 \in E_2$, we have $\sigma(v_2) < \sigma(u)$ or $\sigma(v) < \sigma(u_2)$; see the black and the gray version of e_2 in Case 1, fig. 1, respectively. In both cases, e_1 and e_2 do not cross.

Case 2: For every edge $e \in E(G)$, it holds that $|\texttt{span}(e)| \leq m/3$.

Let $M \subseteq E(G)$ be the collection of all maximal edges of G in σ. Let $\mu = |M|$, and let $M = \{(u_1, v_1), \ldots, (u_\mu, v_\mu)\}$, where $\sigma(u_1) < \sigma(u_2) < \cdots < \sigma(u_\mu)$. Note that $|\texttt{left}(u_1)| \leq \cdots \leq |\texttt{left}(u_\mu)|$ and that $|\texttt{left}(u_1)| = 0$. The equality is due to the fact that u_1 is the first non-isolated vertex of G in σ (and v_1 is the rightmost neighbor of u_1).

Let $a \in [p]$ be the largest index such that $|\texttt{left}(u_a)| \leq m/3$. Since $|\texttt{left}(u_1)| = 0$, it is clear that such an index a exists.

We claim that $m/3 < |\texttt{left}(u_{a+1})| \leq 2m/3 + q$. From the choice of a, it is clear that $|\texttt{left}(u_{a+1})| > m/3$. Note that $\texttt{left}(u_{a+1}) \subseteq \texttt{left}(u_a) \cup \texttt{span}((u_a, v_a)) \cup \texttt{cross}((u_a, v_a))$; see Case 2, fig. 1. This yields our claim since $|\texttt{left}(u_{a+1})| \leq |\texttt{left}(u_a)| + |\texttt{span}((u_a, v_a))| + |\texttt{cross}((u_a, v_a))| \leq 2m/3 + q$.

Now let $X = \texttt{cross}((u_{a+1}, v_{a+1})) \cup \{(u_{a+1}, v_{a+1})\}$, $E_1 = \texttt{left}(u_{a+1})$ and $E_2 = E(G - X) \setminus E_1$. Since $m/3 \leq |\texttt{left}(u_{a+1})| \leq 2m/3 + p$, $|E_1| \leq 2m/3 + p$, and $|E_2| \leq 2m/3$. Finally, we simply move p edges from E_1 to X. Then $|X| \leq 2p + 1$ and $|E_1| \leq 2m/3$. Given our construction, it is clear that no edge in E_1 crosses any edge in E_2; see Case 2, fig. 1.

Case 3: There exists an edge $e^* \in E(G)$ such that $|\texttt{span}(e^*)| > 2m/3$.

Let $e = (u, v)$ be an edge of G such that $|\texttt{span}(e)| > 2m/3$, and there is no $e' \in |\texttt{span}(e)|$ such that $|\texttt{span}(e')| > 2m/3$. Since Case 1 does not apply, for each $e' \in \texttt{span}(e)$, we have $|\texttt{span}(e')| \leq m/3$. Let $V_{uv} = \{w \in V(G) \colon \sigma(u) \leq \sigma(w) \leq \sigma(v)\}$. Let $G_{uv} = G[V_{uv}]$, and let σ_{uv} be the restriction of σ to V_{uv}.

Now we apply Case 2 to the ordered graph (G_{uv}, σ_{uv}). This yields a set $X_{uv} \subseteq E(G_{uv})$ of size at most $q + 1$, and edge sets E_1^{uv} and E_2^{uv} such that $E(G_{uv}) \setminus X_{uv} = E_1^{uv} \uplus E_2^{uv}$, $m/3 \leq |E_1^{uv}| \leq 2m/3$, $|E_2^{uv}| \leq 2m/3$, and no edge in E_1^{uv} crosses any edge in E_2^{uv}.

Let $X = X_{uv} \cup \texttt{cross}(e) \cup \{e\}$. Then $|X| \leq 2(q + 1)$. Let $E_1 = E_1^{uv}$ and $E_2 = E(G - X) \setminus E_1^{uv}$. Since $m/3 \leq |E_1^{uv}| \leq 2m/3$, clearly $|E_2| \leq 2m/3$. It remains to show that no edge of E_2 crosses any edge of E_1; see Case 3, fig. 1. By construction, no edge of E_2^{uv} crosses any edge of E_1^{uv}. The edges in $E_2 \setminus E_2^{uv}$ neither cross e nor do they lie in $\texttt{span}(e)$, so they cannot cross any edge in $E_1 \subseteq \texttt{span}(e)$. ◀

Proof of Theorem 1. We now need to establish a relation between the treewidth of the graph and the size of a balanced separator in it. For this we use the result of Dvořák and Norin [6] that shows a linear dependence between the treewidth and the *separation number* of a graph: the separation number of a graph is the smallest integer s such that every subgraph of the given graph has a balanced separator of size at most s. In other words, they show that if the separation number of the graph is s, then the treewidth of such a graph is $\mathcal{O}(s)$.

Recall that (G, σ, k) is an instance of Ordered Graph Edge Deletion to p-Plane. By Lemma 2, if the ordered graph (G, σ) is $(p + \sqrt{k})$-plane, then the conflict graph $H_{(G,\sigma)}$ has a balanced separator of size at most $2(p + \sqrt{k} + 1)$. Thus, due to the result of Dvořák and Norin [6], the treewidth of $H_{(G,\sigma)}$ is $\mathcal{O}(p + \sqrt{k})$.

Given a graph with N vertices and treewidth t, one can compute, in $(d + 2)^t \cdot N^{\mathcal{O}(1)}$ time, the smallest set of vertices whose deletion results in a graph of degree at most d [14, Folklore]. Applying this result to the conflict graph $H_{(G,\sigma)}$, which has at most $|V(G)|^2 = n^2$ vertices and treewidth $\mathcal{O}(p + \sqrt{k})$, we conclude that Ordered Graph Edge Deletion to p-Plane can be solved in $2^{\mathcal{O}((p+\sqrt{k}) \log p)} \cdot n^{\mathcal{O}(1)}$ time if the given ordered graph (G, σ) is $(p + \sqrt{k})$-plane.

From Branching, we can assume, at the expense of a multiplicative factor of $2^{\mathcal{O}(p\sqrt{k} \log(k+p))}$ on the running time, that the given ordered graph (G, σ) is $(p + \sqrt{k})$-plane. Thus, given (G, σ, k), we can solve Ordered Graph Edge Deletion to p-Plane in $2^{\mathcal{O}(p\sqrt{k} \log(p+k))}$ time. This concludes the proof of Theorem 1. ◀

Open Problems

1. Could we use the concept of the conflict graph for other crossing reduction problems?
2. Is Ordered Graph Edge Deletion to p-Plane $W[1]$-hard with respect to the natural parameter k if p is part of the input? Can we reduce from Independent Set? Note that Vertex Deletion to Degree-p is $W[1]$-hard with respect to treewidth [1] and that outer-p planar graphs have treewidth $O(p)$ [3] (which also follows from theorem 2).
3. What if the vertex order is not given? In other words, what is the parametrized complexity of edge deletion to outer-p planarity?
4. What about computing the crossing number of an ordered graph? Of course, the total number of crossings is always at least the number of crossings per edge.

References

1. N. Betzler, R. Bredereck, R. Niedermeier, and J. Uhlmann. On bounded-degree vertex deletion parameterized by treewidth. *Discrete Appl. Math.*, 160(1):53–60, 2012. doi:j.dam.2011.08.013.
2. S. Cabello and B. Mohar. Adding one edge to planar graphs makes crossing number and 1-planarity hard. *SIAM J. Comput.*, 42(5):1803–1829, 2013. doi:10.1137/120872310.
3. S. Chaplick, M. Kryven, G. Liotta, A. Löffler, and A. Wolff. Beyond outerplanarity. In F. Frati and K.-L. Ma, editors, *Proc. 25th Int. Symp. Graph Drawing & Network Vis. (GD'17)*, volume 10692 of *LNCS*, pages 546–559. Springer, 2018. URL: https://arxiv.org/abs/1708.08723, doi:10.1007/978-3-319-73915-1_42.
4. H. Donkers, B. M. P. Jansen, and M. Włodarczyk. Preprocessing for outerplanar vertex deletion: An elementary kernel of quartic size. *Algorithmica*, 84(11):3407–3458, 2022. doi:10.1007/s00453-022-00984-2.
5. V. Dujmović and S. Whitesides. An efficient fixed parameter tractable algorithm for 1-sided crossing minimization. *Algorithmica*, 40(1):15–31, 2004. doi:10.1007/s00453-004-1093-2.
6. Z. Dvořák and S. Norin. Treewidth of graphs with balanced separations. *J. Comb. Theory, Ser. B*, 137:137–144, 2019. doi:10.1016/j.jctb.2018.12.007.
7. M. R. Fellows, J. Guo, H. Moser, and R. Niedermeier. A generalization of Nemhauser and Trotter's local optimization theorem. *J. Comput. Syst. Sci.*, 77(6):1141–1158, 2011. doi:10.1016/j.jcss.2010.12.001.
8. M. Grohe. Computing crossing numbers in quadratic time. *J. Comput. Syst. Sci.*, 68(2):285–302, 2004. doi:10.1016/j.jcss.2003.07.008.

9 B. M. P. Jansen, D. Lokshtanov, and S. Saurabh. A near-optimal planarization algorithm. In C. Chekuri, editor, *Proc. Ann. ACM-SIAM Symp. Discrete Algorithms (SODA)*, pages 1802–1811, 2014. doi:10.1137/1.9781611973402.130.

10 B. M. P. Jansen and M. Włodarczyk. Lossy planarization: a constant-factor approximate kernelization for planar vertex deletion. In S. Leonardi and A. Gupta, editors, *Proc. Ann. ACM SIGACT Symp. Theory Comput. (STOC)*, pages 900–913, 2022. doi:10.1145/3519935.3520021.

11 K. Kawarabayashi. Planarity allowing few error vertices in linear time. In *Proc. Ann. IEEE Symp. Foundat. Comput. Sci. (FOCS)*, pages 639–648, 2009. doi:10.1109/FOCS.2009.45.

12 K. Kawarabayashi and B. A. Reed. Computing crossing number in linear time. In D. S. Johnson and U. Feige, editors, *Proc. 39th Ann. ACM Symp. Theory Comput. (STOC)*, pages 382–390, 2007. doi:10.1145/1250790.1250848.

13 Y. Kobayashi and H. Tamaki. A faster fixed parameter algorithm for two-layer crossing minimization. *Inform. Process. Lett.*, 116(9):547–549, 2016. doi:j.ipl.2016.04.012.

14 M. Lampis and M. Vasilakis. Structural parameterizations for two bounded degree problems revisited. *CoRR*, abs/2304.14724, 2023. doi:10.48550/arXiv.2304.14724.

15 D. Marx and I. Schlotter. Obtaining a planar graph by vertex deletion. *Algorithmica*, 62(3-4):807–822, 2012. doi:10.1007/s00453-010-9484-z.

16 N. Nishimura, P. Ragde, and D. M. Thilikos. Fast fixed-parameter tractable algorithms for nontrivial generalizations of vertex cover. *Discret. Appl. Math.*, 152(1-3):229–245, 2005. doi:10.1016/j.dam.2005.02.029.

17 K. Sugiyama, S. Tagawa, and M. Toda. Methods for visual understanding of hierarchical system structures. *IEEE Trans. Syst. Man Cybernetics*, 11(2):109–125, 1981. doi:10.1109/TSMC.1981.4308636.

18 M. Xiao. On a generalization of Nemhauser and Trotter's local optimization theorem. *J. Comput. Syst. Sci.*, 84:97–106, 2017. doi:10.1016/j.jcss.2016.08.003.

5.2 Clustered Planarity

Giordano Da Lozzo (University of Rome III, IT), Robert Ganian (TU Wien, AT), Siddharth Gupta (University of Warwick – Coventry, GB), Bojan Mohar (Simon Fraser University – Burnaby, CA), Sebastian Ordyniak (University of Leeds, GB), and Meirav Zehavi (Ben Gurion University – Beer Sheva, IL)

A *flat clustered graph* (for short *flat c-graph*) is a pair $(G, \mathcal{V})$ where G is a graph and $\mathcal{V}$ is a partition of the vertices of G. The parts of $\mathcal{V}$ are called *clusters*. A flat c-graph is *independent* if each of its clusters induces an independent set. Given a family $\mathcal{F}$ of connected graphs and an independent flat c-graph $\mathcal{C} = (G, \mathcal{V} = \{V_1, \ldots, V_k\})$, the Planar $\mathcal{F}$-augmentation problem for $\mathcal{C}$ asks to test for the existence of sets Z_i of edges joining pairs of vertices in V_i for $i = 1, \ldots, k$, such that $G \cup \bigcup_{i=1}^{k} Z_i$ is planar and $G(V_i, Z_i) \in \mathcal{F}$ for $i = 1, \ldots, k$. If $\mathcal{F}$ is the family of all paths, then the problem coincides with the Clustered Planarity with Linear Saturators problem, which is known to be NP-complete [1]. On the other hand, if $\mathcal{F}$ is the family of all trees, then the problem coincides with the Clustered Planarity problem, which has been recently shown to be solvable in polynomial time [2].

We consider Planar $\mathcal{F}$-augmentation for two classes $\mathcal{F}$, i.e., the set $\mathcal{P}$ of all paths and the set $\mathcal{S}$ of all stars. Since it is known that Planar $\mathcal{P}$-augmentation is NP-hard even if the input graph has no edges, but has many non-trivial clusters (clusters containing more than one vertex), we tried to obtain an fpt-algorithm for the parameter treewidth plus the number of non-trivial clusters using SPQR-trees. Unfortunately, after some consideration, we had to abandon this approach, after realizing that we were not able to handle spiralling paths in the solution; see also Figure 4b) in [3] giving a worst-case example for spiralling paths.

We then focused on obtaining exact algorithms for PLANAR $\mathcal{P}$-AUGMENTATION and we are now pretty confident that we can obtain a single-expoential $2^{O(n)}$ algorithm for the problem in the variable embedding case as well as a subexponential $2^{O(\log n\sqrt{n})}$ algorithm in the fixed embedding case (if the initial graph is connected). Both algorithms start by guessing a separator of the solution graph. Then, in the variable embedding case we additionally guess the partition of the separation (which dominates the run-time in this case and is the reason that we do not obtain a subexponential-time algorithm). After doing so the algorithm proceeds recursively in both parts. In the fixed embedding and connected case, we can circumvent guessing the partition, by guessing the curve through the separator, which leaves only $n^{\sqrt{n}} = 2^{O(\log n\sqrt{n})}$ many choices. We leave it open whether either the variable-embedding case can be improved to a subexponential-time algorithm and whether we can avoid the $\log n$ factor in the exponent for the fixed-embedding case.

Our second main result concerns the PLANAR $\mathcal{S}$-AUGMENTATION problem. Here, we can show that if the input graph is bi-connected, then PLANAR $\mathcal{S}$-AUGMENTATION can be solved in polynomial-time in the variable-embedding case. The main ideas behind this algorithm are as follows. Consider the graph H obtained from G after contradicting every cluster into a single vertex such that every half-edge in H is assigned the color that is equal to the vertex in the cluster that it originated from. Our first observation is that G has a solution if and only if H has a drawing such that all but one color of the half-edges around every vertex is consecutive. In other words there is at most one color that is split at every vertex. We now employ a dynamic programming algorithm along an SPQR-tree of H to decide whether H admits such a drawing. To do so we define the following records for every pertinent graph corresponding to an artificial arc st: a record stores the colors of the two outermost half-edges incident to s, the colors of the two outermost half-edges incident to t, and the color that is split (if there is no split color this remains undefined). It is important to note that every pertinent graph st has at most 3 colors that also occur outside (otherwise at least two colors are split and there is no solution for H) and that we do not need to know the identity of any color that does not occur outside. Therefore, the number of possible colors that we store for every entry of the record is at most 4, i.e., the color could be one of the at most 3 colors that also occur outside or it could be another color (which we will denote with the symbol u). Therefore, the number of records for a pertinent graph is at most $4^4 + 5$ (4 options for the colors of the outer half-edges and 5 options for the split color). It is then relatively straightforward to show how the records are computed for S-nodes. For P-nodes, we design a slightly more involved greedy algorithm that guesses the colors of the outer half-edges and then proceeds by computing the forced implications. Rather more complicated is the algorithm for computing R-nodes. Here, we manage to obtain an encoding into 2-SAT after some preprocessing of the instance. It is not clear to us how to get around the condition that G has to be bi-connected. Neither do we know, whether a similar approach can be used for the fixed-embedding case. The main problem with using our approach for the fixed embedding case is that our initial observation about contradicting every cluster into one vertex does not longer preserve solutions. This is because the color that is split at any vertex depends on, where the contracted vertex is placed within the drawing.

References

1 Patrizio Angelini, Giordano Da Lozzo, Giuseppe Di Battista, Fabrizio Frati, Maurizio Patrignani, Ignaz Rutter: *Intersection-Link Representations of Graphs*. J. Graph Algorithms Appl. 21(4): 731-755 (2017)

2 Radoslav Fulek, Csaba D. Tóth: *Atomic Embeddability, Clustered Planarity, and Thickenability*. J. ACM 69(2): 13:1-13:34 (2022)

3 Patrizio Angelini, Giordano Da Lozzo, Giuseppe Di Battista, Fabrizio Frati, Maurizio Patrignani: *2-Level Quasi-Planarity or How Caterpillars Climb (SPQR-)Trees*. CoRR: abs/2011.02431 (2020)

5.3 The Maximum Bimodal Subgraph Problem: Parameterized Complexity and Approximation Scheme

Walter Didimo (University of Perugia, IT), Fedor V. Fomin (University of Bergen, NO), Petr A. Golovach (University of Bergen, NO), Tanmay Inamdar (University of Bergen, NO), Stephen G. Kobourov (University of Arizona – Tucson, US), and Marie Diana Sieper (Universität Würzburg, DE)

A *plane digraph* G is a directed planar graph with a given planar embedding. In particular, the embedding of G defines the circular ordering of the edges incident to each vertex. A vertex v of G is *bimodal* if all its incoming edges (and hence all its outgoing edges) are consecutive in the cyclic order around v. A plane digraph is *bimodal* if all its vertices are bimodal. For example, the plane digraph of Figure 2(a) is not bimodal, namely the vertices $3, 4, 5, 7$ are bimodal while the vertices $1, 2, 6$ are not.

Bimodality is a central property for many graph drawing styles. In particular, it is a necessary condition for the existence of *level-planar* and *upward planar* drawings, where the edges are represented as curves monotonically increasing in the upward direction according to their orientations [11, 12, 13, 18]. Bimodality becomes also a sufficient condition for the existence of embedding-preserving *quasi-upward planar* drawings, in which edges are allowed to violate the upward monotonicity a finite number of times at points called *bends* [5, 6, 7]. Recently, it has been shown that bimodality is also sufficient for the existence of *planar L-drawings* of digraphs, in which distinct L-shaped edges may overlap but not cross [1, 2, 3].

Given a digraph G it is possible to decide in linear time if G admits a bimodal planar embedding, by simply transforming the problem into a planarity testing problem on a suitable digraph obtained by splitting each non-bimodal vertex of G into two vertices [5]. However, when G is not bimodal, a problem that naturally arises is to extract from G a subgraph of maximum size (i.e., with the maximum number of edges) that is bimodal. Unfortunately, this problem is known to be NP-hard, even when G has a given planar embedding and we look for an embedding-preserving maximum bimodal subgraph [8]. This problem is called the *Maximum Bimodal Subgraph* (MBS) problem. For example, Figure 2(b) shows a bimodal subgraph of maximum size for the graph of Figure 2(a).

Binucci et al. [8] describe a heuristic and an exponential-time branch-and-bound algorithm to solve MBS, by also addressing the more specific problem of extracting from the input plane digraph G an upward planar digraph of maximum size.

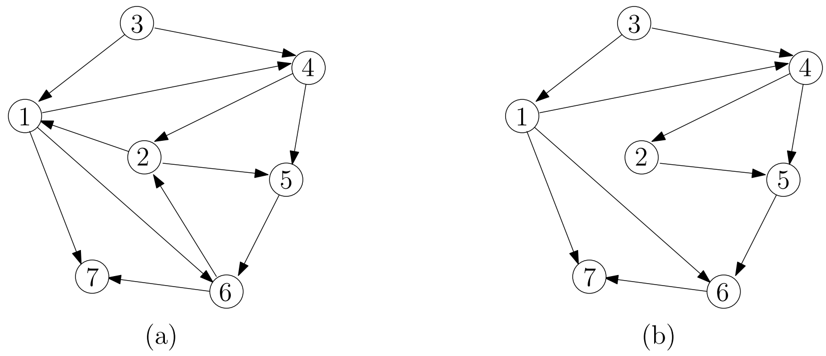

Figure 2 (a) A plane digraph G. (b) A maximum bimodal subgraph of G that preserves its embedding.

Contribution

During the Dagstuhl Seminar we studied the MBS problem from the parameterized complexity and approximability perspectives (we refer to the books [10, 15] for an introduction to the parameterized complexity area). More precisely, we considered the following more general version of the problem with weighted edges, which coincides with MBS if we restrict to unit edge weights.

MWBS(G, w) (*Maximum Weighted Bimodal Subgraph*). *Given a plane digraph G and an edge-weight function $w : E(G) \to \mathbb{Q}^+$, compute a bimodal subgraph of G of maximum weight, i.e., whose sum of the edge weights is maximum over all bimodal subgraphs of G.*

We studied different parameters for the MWBS(G, w) problem and mainly focused on the following: (i) the number b of non-bimodal vertices of the input graph G, which intrinsically captures the nature and the difficulty of the problem; (ii) the branchwidth and the treewidth of the undirected underlying graph of G, which are structural parameters commonly adopted in parameterized complexity. Our main results can be summarized as follows.

- *Structural parameterization.* We show that MWBS is FPT when parameterized by the branchwidth of the input digraph G or, equivalently, by the treewidth of G, which is correlated to the branchwidth by a constant factor [19]. Our algorithm deviates from a standard dynamic approach for graphs of bounded treewidth. Since we have to incorporate the "topological" information about the given embedding in the dynamic program, we exploit the sphere-cut decomposition of Dorn et al. [14].
- *Kernelization.* We prove the existence of a polynomial kernel for the decision version of MWBS parameterized by the number b of non-bimodal vertices. Our kernelization consists of several steps. First we show how to reduce the instance to an equivalent instance whose branchwidth is $O(\sqrt{b})$. Second, by using specific gadgets, we compress the problem to an instance of another problem whose size is bounded by a polynomial of b. Finally, by the standard arguments, [15, Theorem 1.6], based on a polynomial reduction between any NP-complete problems, we obtain a polynomial kernel for MWBS.

  By pipelining the crucial step of the kernelization algorithm with the branchwidth algorithm, we obtain a parameterized subexponential algorithm for MWBS of running time $2^{O(\sqrt{b})} \cdot n^{O(1)}$. Since $b \leq n$, this also implies an algorithm of running time $2^{O(\sqrt{n})}$. We remark that our algorithms are asymptotically optimal up to the *Exponential Time Hypothesis* (ETH) [16, 17]. The NP-hardness result of MBS (and hence of MWBS) given in [8] exploits a reduction from PLANAR-3SAT. The number of non-bimodal vertices in the resulting instance of MBS is linear in the size of the PLANAR-3SAT instance. Using the standard techniques for computational lower bounds for problems on planar graphs [10], we obtain that the existence of an $2^{o(\sqrt{b})} \cdot n^{O(1)}$-time algorithm for MBWS would contradict ETH.

- *Approximability.* We provide an Efficient Polynomial-Time Approximation Scheme (EPTAS) for MWBS, based on Baker's (or shifting) technique [4]. More precisely, using our algorithm for graphs of bounded branchwidth, we give an $(1+\epsilon)$-approximation algorithm that runs in $2^{O(1/\epsilon)} \cdot n^{O(1)}$ time.

Future research

A natural future research direction is to extend our study to a generalization of bimodality, called *k-modality*. Given a positive even integer k, a plane digraph is *k-modal* if the edges at each vertex can be grouped into at most k sets of consecutive edges with the same orientation [20]. For example, it is known that 4-modality is necessary for planar L-drawings [9].

References

1 Patrizio Angelini, Steven Chaplick, Sabine Cornelsen, and Giordano Da Lozzo. Planar L-drawings of bimodal graphs. *J. Graph Algorithms Appl.*, 26(3):307–334, 2022.

2 Patrizio Angelini, Steven Chaplick, Sabine Cornelsen, and Giordano Da Lozzo. On upward-planar l-drawings of graphs. In Stefan Szeider, Robert Ganian, and Alexandra Silva, editors, *47th International Symposium on Mathematical Foundations of Computer Science, MFCS 2022, August 22-26, 2022, Vienna, Austria*, volume 241 of *LIPIcs*, pages 10:1–10:15. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022.

3 Patrizio Angelini, Giordano Da Lozzo, Marco Di Bartolomeo, Valentino Di Donato, Maurizio Patrignani, Vincenzo Roselli, and Ioannis G. Tollis. Algorithms and bounds for L-drawings of directed graphs. *Int. J. Found. Comput. Sci.*, 29(4):461–480, 2018.

4 Brenda S. Baker. Approximation algorithms for NP-complete problems on planar graphs. *J. Assoc. Comput. Mach.*, 41(1):153–180, 1994.

5 Paola Bertolazzi, Giuseppe Di Battista, and Walter Didimo. Quasi-upward planarity. *Algorithmica*, 32(3):474–506, 2002.

6 Carla Binucci, Emilio Di Giacomo, Giuseppe Liotta, and Alessandra Tappini. Quasi-upward planar drawings with minimum curve complexity. In Helen C. Purchase and Ignaz Rutter, editors, *Graph Drawing and Network Visualization – 29th International Symposium, GD 2021, Tübingen, Germany, September 14-17, 2021, Revised Selected Papers*, volume 12868 of *Lecture Notes in Computer Science*, pages 195–209. Springer, 2021.

7 Carla Binucci and Walter Didimo. Computing quasi-upward planar drawings of mixed graphs. *Comput. J.*, 59(1):133–150, 2016.

8 Carla Binucci, Walter Didimo, and Francesco Giordano. Maximum upward planar subgraphs of embedded planar digraphs. *Comput. Geom.*, 41(3):230–246, 2008.

9 Steven Chaplick, Markus Chimani, Sabine Cornelsen, Giordano Da Lozzo, Martin Nöllenburg, Maurizio Patrignani, Ioannis G. Tollis, and Alexander Wolff. Planar l-drawings of directed graphs. In Fabrizio Frati and Kwan-Liu Ma, editors, *Graph Drawing and Network Visualization – 25th International Symposium, GD 2017, Boston, MA, USA, September 25-27, 2017, Revised Selected Papers*, volume 10692 of *Lecture Notes in Computer Science*, pages 465–478. Springer, 2017.

10 Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms.* Springer, 2015.

11 Giuseppe Di Battista, Peter Eades, Roberto Tamassia, and Ioannis G. Tollis. *Graph Drawing: Algorithms for the Visualization of Graphs.* Prentice-Hall, 1999.
12 Giuseppe Di Battista and Enrico Nardelli. Hierarchies and planarity theory. *IEEE Trans. Syst. Man Cybern.*, 18(6):1035–1046, 1988.
13 Walter Didimo. Upward graph drawing. In *Encyclopedia of Algorithms*, pages 2308–2312. 2016.
14 Frederic Dorn, Eelko Penninkx, Hans L. Bodlaender, and Fedor V. Fomin. Efficient exact algorithms on planar graphs: Exploiting sphere cut decompositions. *Algorithmica*, 58(3):790–810, 2010.
15 Fedor V. Fomin, Daniel Lokshtanov, Saket Saurabh, and Meirav Zehavi. *Kernelization. Theory of parameterized preprocessing.* Cambridge University Press, Cambridge, 2019.
16 Russell Impagliazzo and Ramamohan Paturi. Complexity of k-sat. In *Proceedings of the 14th Annual IEEE Conference on Computational Complexity, Atlanta, Georgia, USA, May 4-6, 1999*, pages 237–240. IEEE Computer Society, 1999.
17 Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *J. Comput. Syst. Sci.*, 63(4):512–530, 2001.
18 Michael Jünger, Sebastian Leipert, and Petra Mutzel. Level planarity testing in linear time. In Sue Whitesides, editor, *Graph Drawing, 6th International Symposium, GD'98, Montréal, Canada, August 1998, Proceedings*, volume 1547 of *Lecture Notes in Computer Science*, pages 224–237. Springer, 1998.
19 Neil Robertson and Paul D. Seymour. Graph minors. X. obstructions to tree-decomposition. *J. Comb. Theory, Ser. B*, 52(2):153–190, 1991.
20 Juan José Besa Vial, Giordano Da Lozzo, and Michael T. Goodrich. Computing k-modal embeddings of planar digraphs. In Michael A. Bender, Ola Svensson, and Grzegorz Herman, editors, *27th Annual European Symposium on Algorithms, ESA 2019, September 9-11, 2019, Munich/Garching, Germany*, volume 144 of *LIPIcs*, pages 19:1–19:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019.

5.4 Parameterised Algorithms for Sequential Crossing Number

Eduard Eiben (Royal Holloway, University of London, GB), Thekla Hamm (Utrecht University, NL), Petr Hliněný (Masaryk University – Brno, CZ), Martin Nöllenburg (TU Wien, AT), Ignaz Rutter (Universität Passau, DE), and Manuel Sorge (TU Wien, AT)

A driving motivation of the field of graph drawing is the visualisation of information that can be captured by graphs. Often such information can evolve over time and hence the treatment of temporal aspects is an important challenge in graph drawing. In this context it makes sense to require some "stability" of the drawing of a graph that changes over time in the sense that parts of the graph continue to occur in consecutive time steps are not redrawn in a different way. Motivated by this we define the following crossing-minimal drawing variant. Below $\mathrm{cr}(\mathcal{G})$ for a drawing $\mathcal{G}$ of a graph denotes the number of edge crossings in the drawing.

SEQUENTIAL CROSSING NUMBER (SCN):
Problem: A sequence $G_1, G_2, \ldots, G_t$ of (in general non-disjoint) graphs, an integer k.
Question: Is there a drawing $\mathcal{G}_i$ of each G_i such that for each $i \in [t-1]$ we have

$$\mathcal{G}_i|_{G_i \cap G_{i+1}} = \mathcal{G}_{i+1}|_{G_i \cap G_{i+1}} \text{ and } \sum_{i \in [t]} \mathrm{cr}(\mathcal{G}_i) \leq k?$$

Relations to other problems and implied hardness results

SCN is obviously more general than the classical crossing number problem, which corresponds to the setting in which $t = 1$, and it is also more general than the partially predrawn crossing number problem [5] which can be captured using $t = 2$, having as G_1 the rigidised (also adding many parallel edges) prescribed partial drawing of the graph and as G_2 the graph to which the partial predrawing should be extended.

Moreover, in the special case that for all $i, j \in [t]$, $G_i \cap G_j = \bigcap_{\ell \in [t]} G_i$, i.e. the pairwise intersection of all given graphs is the same, then we speak of *sunflower instances*, and refer to $\bigcap_{\ell \in [t]} G_i$ as the *core* (in line with terminology used for Simultaneous Embedding with Fixed Edges (SEFE) [2]). On sunflower instances, SEFE and SCN with $k = 0$ are easily seen to be completely equivalent.

Indeed, there is even a simple reduction from SEFE on general (not necessarily sunflower) instances to SCN with $k = 0$ by considering, for an input $G_1, \ldots, G_t$ of SEFE, the instance of SCN (on sunflower instances) given by $G_\ell \cup \bigcup_{i,j \in [t]} G_i \cap G_j$, for $\ell \in [t]$, and $k = 0$.

From the known hardness of SEFE on sunflower instances with $t = 3$ and the core being a star[1], we obtain the following.

▶ **Theorem 1.** *SCN is* NP*-hard, even on sunflower instances with the core being a star,* $t = 3$ *and* $k = 0$*.*

A generalisation of SEFE to allow for drawings with crossings in which the total number of crossings over any single graph is considered as the decision value has been proposed as Simultaneous Crossing Number [3], and can be reduced to SCN (with general k) in the same way. Of course, since both problems are NP-complete, there is also a polynomial reduction from SCN to the simultaneous crossing number problem. However, a difference between both reductions is that, in the direction from Simultaneous Crossing Number to SCN, we never increase the structural complexity of the union of all instance graphs, whereas this is necessary in the direction from SCN to Simultaneous Crossing Number to allow for 'reappearing' parts of graphs to be drawn differently.

SCN with $k = 0$ also generalises planarity of streamed graphs [4]. A *streamed graph* with *time window* $w \in \mathbb{N}$ and backbone is given by a vertex set V, a set of edges B on V called a *backbone*, and an ordered non-repeating sequence of edges $e_1, \ldots, e_m$ on V which are not in B. By definition, this streamed graph is planar if and only if the following is a yes-instance for SCN with $k = 0$: For $t = m$ and for $i \in [t]$, $G_i := (V, B \cup \{e_j \mid i \leq j < i + w\})$.

From the hardness result for testing planarity of streamed graphs with time window 5 and empty backbones [4] we can infer the following for SCN.

▶ **Theorem 2.** *SCN is* NP*-hard, even on instances with* $|E(G_i)| \leq 5$ *for all* $i \in [t]$ *and* $k = 0$*.*

Parameterised Algorithms for Sequential Crossing Number

In view of the above hardness results, our main target during the seminar was to formulate parameterised algorithms for SCN. Notably, we believe to be able to give the following results.

1. SCN parameterised by $k + \max_{i \in [t-1]} |V(G_i) \cap V(G_{i+1})|$ is in FPT.
2. SCN parameterised by k is in XP on inputs in which for all $i \in [t-1]$, $G_i \cap G_{i+1}$ is 3-connected. (This is subsumed by the next FPT result but also used in its proof with $k = 0$.)

3. SCN parameterised by k is in FPT on inputs in which for all $i \in [t-1]$, $G_i \cap G_{i+1}$ is 3-connected.
4. SCN parameterised by $k + t + r$ is in FPT, where r is an upper bound on the maximum number of bridges of a P-node in the SPQR-tree and the maximum degree of a vertex in the block cut tree of the intersection of two consecutive input graphs, on instances in which for all $i \in [t-1]$, $G_i \cap G_{i+1}$ is connected.
   - Notice that for 3-connected consecutive intersections, $r = 0$ and the condition is obviously satisfied. The only reason that this result does not subsume the FPT-result for 3-connected consecutive intersections is the additional parameterisation on t.
   - The hardness from Theorem 1 implies that we cannot hope to drop r – the maximum degree of a vertex in the block cut tree of the intersection of two consecutive input graphs, from the problem parameter.

In terms of techniques, the first two results are achieved by dynamic programming and branching algorithms, respectively, the former of which we briefly sketch:

▶ **Theorem 3.** *SCN parameterised by* $k + \max_{i\in[t-1]} |V(G_i) \cap V(G_{i+1})|$ *is in* FPT*.*

Idea. We carry out dynamic programming on t, where the dynamic programming table is indexed by pairs of a drawing of $G_{i-1} \cap G_i$ and a drawing of $G_i \cap G_{i+1}$ with at most k crossings in total. The number of such pairs of drawings is easily seen to be bounded in the considered parameter. Each table entry stores the minimum total number of crossings so far (using the state from the previous dynamic programming step) under the condition that the drawing of G_i coincides with the index-prescribed drawings of $G_{i-1} \cap G_i$ and $G_i \cap G_{i+1}$. To compute an entry for a table index we use the FPT-algorithm for partially predrawn crossing number [5]. ◀

The second two results are far more involved: Similarly to the fixed-parameter algorithm we use a bicriteria approach to reduce the treewidth of the input graphs to a function of the considered parameters – and, more importantly, the treewidth of a derived auxiliary graph that arises from the disjoint union of all G_i by identifying vertices in the intersections (both edges and vertices) of consecutive G_i with one another. For Result 3 this needs to be preceded by a careful reduction of t to a function of k using Result 2. Once we are in a situation of bounded treewidth, we want to invoke Courcelle's theorem on a bounded-length MSO-formula that encodes the fact that the considered auxiliary graph has a drawing with at most k-crossings. However, previously no such formula was known. To overcome this difficulty, we formulate a result of independent interest:

- We give an $f(k)$-length MSO-formula for expressing the existence of a drawing of a connected graph G with at most k crossings on a graph derived in polynomial time from G in a way that crossings and rotations around cut vertices of P-nodes and high-degree cut vertices are given by the interpretation of the free variables.

The crucial advantage over the known formulations for expressing the existence of a drawing of a connected graph G with at most k crossings via excluding Kuratowski minors after identifying crossings (which is also used in the known FPT-algorithm for Crossing Number) is that the known formulations give us no direct knowledge of the drawing. Let us highlight the fact that the graph derived from G has treewidth bounded in $\mathrm{tw}(G) + r$, which is where r comes in as a parameter in Result 4.

References

1. P. Angelini, G. Da Lozzo and D. Neuwirth, *Advancements on SEFE and Partitioned Book Embedding problems*, Theoretical Computer Science, 2015, pp. 71-89, doi: 10.1016/j.tcs.2014.11.016.
2. T. Bläsius, S. G. Kobourov and I. Rutter, *Simultaneous Embedding of Planar Graphs*, Handbook on Graph Drawing and Visualization, 2013, pp.349-381.
3. M. Chimani, M. Jünger and M. Schulz, *Crossing Minimization meets Simultaneous Drawing*, 2008 IEEE Pacific Visualization Symposium, 2008, pp. 33-40, doi: 10.1109/PACIFICVIS.2008.4475456.
4. G. Da Lozzo and I. Rutter, *Planarity of streamed graphs*, Theoretical Computer Science, 2019, pp. 1-21, doi: 10.1016/j.tcs.2019.09.029.
5. T. Hamm and P. Hliněný, *Parameterised Partially-Predrawn Crossing Number*, 38th International Symposium on Computational Geometry (SoCG 2022), 2022, doi: 10.4230/LIPIcs.SoCG.2022.46.

5.5 Upward and Rectilinear Planarity Testing and *st*-Planar Edge Completion

Bart M. P. Jansen (TU Eindhoven, NL), Liana Khazaliya (TU Wien, AT), Philipp Kindermann (Universität Trier, DE), Giuseppe Liotta (University of Perugia, IT), Fabrizio Montecchiani (University of Perugia, IT), Kirill Simonov (Hasso Plattner Institute, University of Potsdam, DE)

The first problems considered by this working group are Upward and Rectilinear Planarity Testing. Given a directed acyclic graph G, upward planarity testing asks whether G admits a crossing-free drawing where all edges are monotonically increasing in a common direction, which is conventionally called the upward direction; see Fig.3. For an undirected graph G, rectilinear planarity testing asks whether G admits a crossing-free drawing such that each edge is either a horizontal or a vertical segment; see Fig.3. Both upward planarity and rectilinear planarity testing are classical and extensively investigated topics in graph drawing (see, e.g. [1, 19, 15, 18]). It is known that they are both NP-complete in the so-called *variable embedding setting*, that is when the testing algorithm must verify whether the input graph has a combinatorial structure of its faces which allows the balancing of angles described above. Again unsurprisingly, both proofs of NP-completeness follow the same logic based on a reduction from a common flow problem on planar graphs [20]. These NP-completeness results have motivated a flourishing literature describing both polynomial-time solutions for special classes of graphs and parameterized solutions for general graphs. For example, polynomial-time solutions are known for both problems when the input graph has treewidth at most two [10, 12, 13, 6, 11]; also, rectilinear planarity testing can be solved in linear time if the maximum degree of the input graph is at most three [14], and upward planarity testing can be solved in linear time if the digraph has only one source vertex [2, 4, 21]. Concerning parameterized solutions, upward planarity testing is fixed-parameter tractable when parameterized by the number of triconnected components [5], by the treedepth [7], and by the number of sources [7].

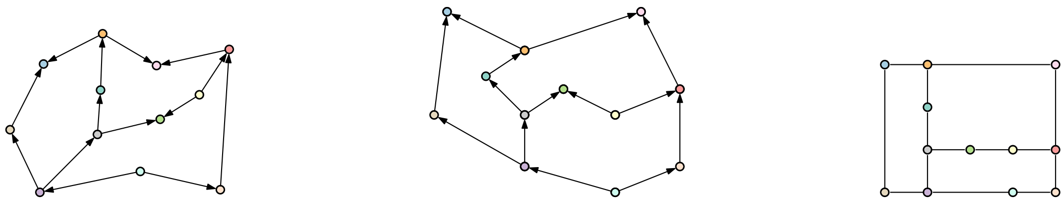

Figure 3 A directed graph $\overrightarrow{G}$, an upward planar drawing of $\overrightarrow{G}$, and a rectilinear planar drawing of its underlying undirected graph G.

The research in this paper is motivated by the fact that both upward planarity and rectilinear planarity testing are known to lie in XP when parameterized by treewidth [7, 11]. Determining whether these two parameterized problems are in FPT are mentioned as open problems in both [7, 11]. The main contribution of this paper is as follows.

▶ **Theorem 1.** *Upward planarity testing and rectilinear planarity testing parameterized by treewidth are both W[1]-hard. Moreover, assuming the Exponential Time Hypothesis, neither problem can be solved in time $f(k) \cdot n^{o(k)}$ for any computable function f, where k is the treewidth of the input graph.*

Fig.1 implies that, under the standard hypothesis FPT $\neq$ W[1] in parameterized complexity, there exists no fixed-parameter tractable algorithm for either problem parameterized by treewidth, hence answering the above mentioned open problems. To obtain our results we analyze the auxiliary flow problem used as a common starting point in the NP-completeness proof of both planarity problems. It closely resembles the All-or-Nothing Flow problem (AoNF), which asks for an st-flow of prescribed value in an edge-capacitated flow network such that each edge is either used fully, or not at all. The AoNF problem parameterized by treewidth was recently shown to be W[1]-hard (in fact, even XNLP-complete) on general graphs by Bodlaender et al. [3]. By a significant adaptation of their construction, we can prove that AoNF parameterized by treewidth remains W[1]-hard on *planar* graphs. By revisiting the chain of reductions to the planarity testing problems, passing through the Circulating Orientation problem in between, we show they can be carried out without blowing up the treewidth of the graph and thereby obtain Fig.1.

Last but not least, we also considered the related st-Planar Edge Completion Problem. Let $G = (V, E)$ be a digraph. A vertex of G with no incoming edges is a *source* of G, while a vertex without outgoing edges is a *sink* of G. A digraph G is an *st-planar graph* if it admits a planar embedding such that: (1) it contains no directed cycle; (2) it contains a single source vertex s and a single sink vertex t; (3) s and t both belong to the external face of the planar embedding.

Upward planarity, as was also mentioned above, is a rather natural and well-studied notion of planarity for directed graphs (see, e.g., [7, 8, 1, 9, 20, 16]). In particular, a planar digraph is *upward* if it admits a planar drawing where all edges are oriented upward. A well-known result in graph drawing states that a digraph G is upward if and only if G is a subgraph of an st-planar graph [1, 9][2]. However, since testing for upward planarity is an NP-complete problem already for biconnected graphs [20], determining whether a biconnected

[2] From the proof in Lemma 4.1 of [9], one can in fact observe that a digraph is upward planar if and only if it is a subgraph of an st-planar graph defined over the same set of vertices.

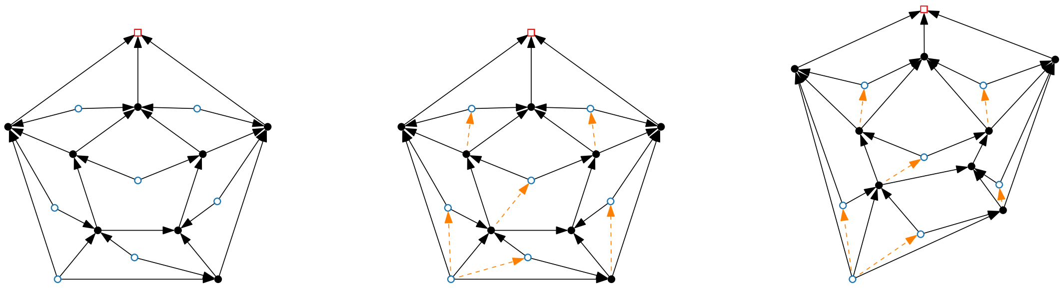

Figure 4 (a) A digraph G with $2k+1=7$ sources and 1 sink; G has a unique planar embedding up to the choice of the external face; (b) A completion of G to an st-planar graph obtained by adding $2k=6$ edges; (c) An upward planar drawing of the completion of G.

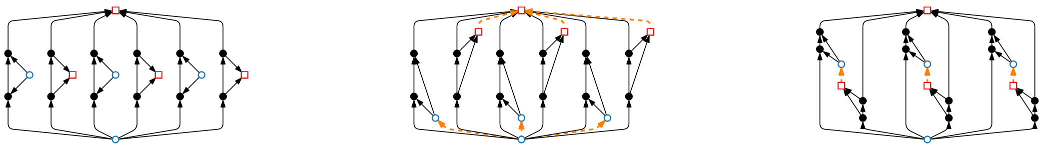

Figure 5 (a) A biconnected digraph G with 4 sources and 4 sinks; (b) With the given embedding, 6 edges have to be added to complete G to an st-planar graph; (c) With a different embedding, adding 3 edges is sufficient.

graph is a subgraph of an st-planar graph is also computationally challenging. On the other hand, checking whether a digraph is st-planar can be done efficiently in polynomial time. This observation motivates for the investigation of the following problem.

> st-Planar Edge Completion (st-PEC)
> **Input:** A biconnected digraph G
> **Parameter:** $k \in \mathbb{N}$
> **Question:** Is it possible to add at most k edges to G such that the resulting graph is an st-planar graph?

We came up with a fixed-parameter tractable algorithm for the st-Planar Edge Completion problem. To help understanding the combinatorial and algorithmic challenges behind the problem, we make the observation that the parameter k provides an upper bound on the number of sources and sinks in the input digraph G. Since an edge can remove the presence of at most one source and one sink, if the total number of sources and sinks in G exceeds $2k+2$, we can promptly reject the instance. Conversely, a positive answer to st-Planar Edge Completion implies that G is upward planar. In this respect, it is worth mentioning that Chaplick et al. [7] have previously demonstrated that testing a digraph for upward planarity is fixed-parameter tractable when parameterized by the number of its sources. However, for every $k \geq 1$, there are upward planar digraphs with at most $2k+1$ sources that cannot be augmented to an st-planar graph by adding k edges; refer to Fig.4 for an illustration. Furthermore, while an upward planarity test halts upon finding an upward planar embedding, not all upward planar embeddings of the same digraph can lead to an st-planar graph after the addition of k edges. Fig.5 demonstrates an upward planar digraph along with two of its upward planar embeddings: the embedding in Fig.5 requires 6 edges to be augmented into an st-planar digraph, whereas the embedding in Fig.5 can be augmented with 3 edges.

In order to overcome the above technical challenges, our result is based on a structural decomposition of the digraph into its triconnected components using SPQR-trees (similarly as done in [7]), as well as on novel insights regarding the combinatorial properties of upward planar digraphs.

References

1 Giuseppe Di Battista, Peter Eades, Roberto Tamassia, Ioannis G. Tollis. Graph Drawing: Algorithms for the Visualization of Graphs. Prentice-Hall, 1999, 0-13-301615-3, https://dblp.org/rec/bib/books/ph/BattistaETT99.

2 Paola Bertolazzi, Giuseppe Di Battista, Carlo Mannino, Roberto Tamassia. Optimal Upward Planarity Testing of Single-Source Digraphs. SIAM J. Comput. https://dblp.org/rec/journals/siamcomp/BertolazziBMT98.

3 Hans L. Bodlaender, Gunther Cornelissen, Marieke van der Wegen. Problems hard for treewidth but easy for stable gonality. Graph-Theoretic Concepts in Computer Science – 48th International Workshop, WG 2022, Tübingen, Germany, June 22-24, 2022, Revised Selected Papers, Lecture Notes in Computer Science, 13453, 84–97, Springer, 2022.

4 Guido Brückner, Markus Himmel, Ignaz Rutter. An SPQR-Tree-Like Embedding Representation for Upward Planarity. Graph Drawing and Network Visualization – 27th International Symposium, GD 2019, Prague, Czech Republic, September 17-20, 2019, Proceedings. https://dblp.org/rec/conf/gd/BrucknerHR19.bib.

5 Hubert Y. Chan, A Parameterized Algorithm for Upward Planarity Testing, Algorithms – ESA 2004, 12th Annual European Symposium, Bergen, Norway, September 14-17, 2004, Proceedings, Lecture Notes in Computer Science,10.1007/978-3-540-30140-0_16.

6 Fabrizio Frati, Planar rectilinear drawings of outerplanar graphs in linear time, Comput. Geom., 10.1016/j.comgeo.2021.101854.

7 Steven Chaplick and Emilio Di Giacomo and Fabrizio Frati and Robert Ganian and Chrysanthi N. Raftopoulou and Kirill Simonov, Parameterized Algorithms for Upward Planarity, 38th International Symposium on Computational Geometry, SoCG 2022, June 7-10, 2022, Berlin, Germany, LIPIcs, 224, 26:1–26:16, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022, 10.4230/LIPIcs.SoCG.2022.26.

8 Steven Chaplick and Emilio Di Giacomo and Fabrizio Frati and Robert Ganian and Chrysanthi N. Raftopoulou and Kirill Simonov, Testing Upward Planarity of Partial 2-Trees, GD, Lecture Notes in Computer Science, 13764, 175–187, Springer, 2022.

9 Giuseppe Di Battista and Roberto Tamassia, Algorithms for Plane Representations of Acyclic Digraphs, Theor. Comput. Sci., 61, 175–198, 1988.

10 Giuseppe Di Battista and Giuseppe Liotta and Francesco Vargiu, Spirality and Optimal Orthogonal Drawings, SIAM J. Comput., 27, 6, 1764–1811, 1998.

11 Emilio Di Giacomo and Giuseppe Liotta and Fabrizio Montecchiani, Orthogonal planarity testing of bounded treewidth graphs, J. Comput. Syst. Sci., 125, 129–148, 2022, 10.1016/j.jcss.2021.11.004.

12 Walter Didimo and Francesco Giordano and Giuseppe Liotta, Upward Spirality and Upward Planarity Testing, SIAM J. Discret. Math., 23, 4, 1842–1899, 2009, 10.1137/070696854.

13 Walter Didimo and Michael Kaufmann and Giuseppe Liotta and Giacomo Ortali, Rectilinear Planarity of Partial 2-Trees, GD, Lecture Notes in Computer Science, 13764, 157–172, Springer, 2022.

14 Walter Didimo and Giuseppe Liotta and Giacomo Ortali and Maurizio Patrignani, Optimal Orthogonal Drawings of Planar 3-Graphs in Linear Time, Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SODA 2020, Salt Lake City, UT, USA, January 5-8, 2020, 806–825, SIAM, 2020, 10.1137/1.9781611975994.49.

15 Takao Nishizeki and Md. Saidur Rahman, Planar Graph Drawing, Lecture Notes Series on Computing, 12, World Scientific, 2004, 10.1142/5648, 981-256-033-5, Mon, 03 Apr 2023 16:01:56 +0200, https://dblp.org/rec/books/ws/NishizekiR04.bib.

16 William T. Trotter and John I. Moore Jr., The dimension of planar posets, J. Comb. Theory, Ser. B, 22, 1, 54–67, 1977.

17 Roberto Tamassia, Handbook on Graph Drawing and Visualization, 1–42, 2013, https://cs.brown.edu/people/rtamassi/gdhandbook/chapters/planarity.pdf.

18 Roberto Tamassia, Handbook on Graph Drawing and Visualization, 1–42, 2013, https://cs.brown.edu/people/rtamassi/gdhandbook/chapters/planarity.pdf.

19 Michael Kaufmann and Dorothea Wagner, Drawing Graphs, Methods and Models (the book grow out of a Dagstuhl Seminar, April 1999), Lecture Notes in Computer Science, 2025, Springer, 2001, 10.1007/3-540-44969-8, https://dblp.org/rec/conf/dagstuhl/1999dg.bib.

20 Ashim Garg and Roberto Tamassia, On the Computational Complexity of Upward and Rectilinear Planarity Testing, SIAM J. Comput., 31, 2, 601–625, 2001.

21 Michael D. Hutton and Anna Lubiw, Upward Planning of Single-Source Acyclic Digraphs, SIAM J. Comput., 25, 2, 291–311, 1996.

Participants

- Akanksha Agrawal
Indian Institute of Techology Madras, IN
- Sergio Cabello
University of Ljubljana, SI
- Giordano Da Lozzo
University of Rome III, IT
- Walter Didimo
University of Perugia, IT
- Eduard Eiben
Royal Holloway, University of London, GB
- Fedor V. Fomin
University of Bergen, NO
- Robert Ganian
TU Wien, AT
- Petr A. Golovach
University of Bergen, NO
- Siddharth Gupta
University of Warwick – Coventry, GB
- Thekla Hamm
Utrecht University, NL
- Petr Hlinený
Masaryk University – Brno, CZ
- Tanmay Inamdar
University of Bergen, NO
- Bart Jansen
TU Eindhoven, NL
- Michael Kaufmann
Universität Tübingen, DE
- Liana Khazaliya
TU Wien, AT
- Philipp Kindermann
Universität Trier, DE
- Stephen G. Kobourov
University of Arizona – Tucson, US
- Giuseppe Liotta
University of Perugia, IT
- Bojan Mohar
Simon Fraser University – Burnaby, CA
- Fabrizio Montecchiani
University of Perugia, IT
- Martin Nöllenburg
TU Wien, AT
- Sebastian Ordyniak
University of Leeds, GB
- Ignaz Rutter
Universität Passau, DE
- Saket Saurabh
The Institute of Mathematical Sciences – Chennai, IN
- Raimund Seidel
Universität des Saarlandes – Saarbrücken, DE
- Roohani Sharma
MPI für Informatik – Saarbrücken, DE
- Marie Diana Sieper
Universität Würzburg, DE
- Kirill Simonov
Hasso-Plattner-Institut, Universität Potsdam, DE
- Manuel Sorge
TU Wien, AT
- Yushi Uno
Osaka Metropolitan University, JP
- Alexander Wolff
Universität Würzburg, DE
- Meirav Zehavi
Ben Gurion University – Beer Sheva, IL

Report from Dagstuhl Seminar 23171

Driving HPC Operations With Holistic Monitoring and Operational Data Analytics

Jim Brandt[*1], **Florina Ciorba**[†2], **Ann Gentile**[†3], **Michael Ott**[†4], and **Torsten Wilde**[†5]

1 **Sandia National Laboratories, US.** `brandt@sandia.gov`
2 **University of Basel, CH.** `florina.ciorba@unibas.ch`
3 **Sandia National Laboratories, US.** `gentile@sandia.gov`
4 **Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities, DE.** `ott@lrz.de`
5 **Hewlett Packard Enterprise – Böblingen, DE.** `wilde@hpe.com`

Abstract

Advances in analytic approaches have brought the vision of efficient High Performance Computing (HPC) operations enabled by dynamic analysis driving automated feedback and adaptation within reach. Many HPC centers have started the development and deployment of frameworks to enable continuous and holistic monitoring, archiving, and analysis of performance data from their production machines and related infrastructures. The impact of such frameworks rests upon the ability to effectively analyze such data and to take action based on analysis results. Analytic techniques have been successfully developed and applied in other domains but their features may not apply directly to HPC operations data and situations. Response options are limited in HPC implementations. Leveraging, adapting, and extending analysis techniques and response options would open up new avenues for research and development of actionable analytics that can drive more intelligent operations through both manual and automated response to conditions of interest.

This Dagstuhl Seminar 23171 brought together practitioners and researchers in the areas of HPC system management and monitoring, analytics, and computer science to collaboratively work on developing community solutions for revolutionizing HPC system operations. The topics discussed in this seminar spanned use cases, data and analytic approaches required to address the use cases, use of analysis results to improve performance and operations, and research in the development and use of autonomous feedback loops.

Seminar April 23–28, 2023 – https://www.dagstuhl.de/23171
2012 ACM Subject Classification General and reference → General literature; Information systems → Information systems applications
Keywords and phrases Monitoring, Operational Data Analytics, Dagstuhl Seminar, WAFVR
Digital Object Identifier 10.4230/DagRep.13.4.98

* Editorial Assistant / Collector
† Editor / Organizer

Driving HPC Operations With Holistic Monitoring and Operational Data Analytics, *Dagstuhl Reports*, Vol. 13, Issue 4, pp. 98–120
Editors: Jim Brandt, Florina Ciorba, Ann Gentile, Michael Ott, and Torsten Wilde
Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Executive Summary

Jim Brandt (Sandia National Laboratories, US, brandt@sandia.gov)
Florina Ciorba (University of Basel, CH, florina.ciorba@unibas.ch)
Ann Gentile (Sandia National Laboratories, US, gentile@sandia.gov)
Michael Ott (Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities, DE, ott@lrz.de)
Torsten Wilde (Hewlett Packard Enterprise, DE, wilde@hpe.com)

The Dagstuhl Seminar 23171 (April 23–28, 2023) brought together 35 practitioners and researchers in the areas of HPC system management and monitoring, data analytics, and computer science to collaboratively work on developing community solutions for revolutionizing HPC system operations. Autonomous operations have long been the vision for efficient HPC system operations due to the size and complexity of current and evolving HPC systems and the need for pervasive, low-latency response. Autonomous operations are a complex topic encompassing monitoring, analysis, feedback, and response. The seminar goals were to make substantial progress on the technical, community, and funding challenges necessary for the community to move forward and reach this vision.

The seminar schedule comprised a mix of keynotes presentations, seed and position talks, enlisted and ad-hoc lightning talks, interleaved with plenary discussions and working group discussions, both in the seminar rooms as well as outdoors.

These program elements and the active participation of the attendees lead to many fruitful discussions on the following topics:

- Center-specific urgent use cases that drive data collection, analysis, and response requirements across the variety of institutions represented.
- Types of available data, including sources, semantics, and fidelity, to support continuous analyses.
- Requirements for actionable analytics. What is needed to convert raw data into information upon which action can be taken (e.g., confidence measures, explainability requirements not inherent in AI approaches, representation of results, latency, etc.)?
- Applicability of existing analytics and informatics approaches to the domain-specifics of HPC operations. While there are many promising ML/AI approaches in other domains (e.g., image/speech processing, autonomous vehicles), it is not yet clear how many and which of those apply to the HPC operations and research domains (e.g., the occurrence of rare fault events, discontinuity of inertia-less measurements).
- Opportunities for response involving infrastructure, hardware, system software, and applications. Identification of feedback hooks that would need to be added to existing and evolving system components (e.g., hardware, firmware, system software, application software) to support automated response.
- Exploration of formalism and architectural design patterns from the field of Self-Adaptive Systems to facilitate common, interoperable, and interchangeable design and development paths forward.

The technical presentations and engaging discussions reinforced the urgent need and desire for a community approach to advance the state and practice of HPC Monitoring and Operational Data Analytics, with the goal of revolutionizing HPC operations and research, in order to deliver efficient and sustainable HPC systems and applications.

The fundamental results of this community discussion are given in Section 10 of this report. These include assessments of the state of autonomous loops in HPC operations and assessments of challenges and opportunities. The community agreed to continue meeting, on a monthly basis and at upcoming community-relevant events. They further agreed to develop *proofs of concepts* for concrete use cases that will showcase both the need for holistic monitoring and analysis and their benefits for more efficient HPC operations. These proofs of concepts will also serve as a basis for technical design decisions and prototype solutions to be deployed in various HPC systems. The final goals of our effort are to continue to build and progress on a community collaborative *technical path* forward, a community *interaction path* forward, and a community *collaborative funding path* forward, as described in Section 11, to fulfill the vision of autonomous and efficient HPC system operations.

2 Table of Contents

3 Keynotes

3.1 Community Readiness and Opportunities for Progress in HPC Monitoring, Analysis, Feedback, and Response

Jim Brandt (Sandia National Laboratories – Livermore and Albuquerque, US, brandt@sandia.gov)

This talk presents work in progress at Sandia National Laboratories. It includes a comprehensive vision on holistic monitoring and feedback as well as the mechanisms currently being used for data collection and feedback. Approaches to Machine Learning that have been explored are presented along with work in progress for turning signals from a ML engine into feedback to drive response for mitigating specific application performance degrading conditions in an HPC Lustre file system.

3.2 Visions for HPC System and Facility onitoring and Operational Data Analytics

Utz-Uwe Haus (HPE HPC/AI EMEA Research Lab – Zurich, CH, utz-uwe.haus@hpe.com)

Starting from the well-known 4x4 table of characteristics for component vs action levels for MODA (Bates et al.) we highlight the importance and tight interconnection to digital twins of data centers, in particular HPC data centers including System level twins. Recent advances in data-aware middlewares, scheduling simulation tools and high frequency monitoring are discussed as essential components in a holistic, sustainability focused approach.

4 Talks: Use Cases

4.1 Continuous Performance Counter Sampling

Michael Ott (Leibniz Supercomputing Centre – Garching, DE, ott@lrz.de)

User applications very often are black boxes for HPC operators in terms of requirements and runtime characteristics. Yet, having this information about user workloads would be beneficial for multiple purposes: procurement of new systems that better fit the user requirements, influencing scheduling decisions to optimize system throughput, or identify jobs with low performance. This talk will present ideas on continuous performance counter sampling to obtain better insight into HPC workloads without user interaction.

4.2 System Monitoring from the Performance Measurement Perspective

Kevin Huck (University of Oregon – Eugene, OR, US, khuck@cs.uoregon.edu)

From the application performance measurement perspective, limited to the user space domain, there are several motivations for runtime monitoring of the application, operating system and hardware. These motivations run the gamut between simple curiosity to identifying the causes of software failures. In this talk, we will explore these motivations and discuss the barriers to getting access to exposed as well as privileged data to contribute to performance understanding.

4.3 HPC Operations and Monitoring

Esa Heiskanen (CSC – IT Center for Science – Kajaani, FI, esa.heiskanen@csc.fi)

Installing and operating large HPC systems is complicated from an operational point of view. There are always unpredictable things that happen and those challenges and risks can be mitigated with good preparation and design of different elements. Having better visibility and predictability to system and applications helps operators to understand the workload of a system and perform preventive anomaly detection and maintenance. Digital twins are defined in 5 levels as descriptive, informative, predictive, comprehensive and autonomous. Building a digital twin of a data center and HPC system is challenging and needs co-operations with multiple parties to achieve goals that give real benefits of digital twin.

5 Position Talks

5.1 Workflows & Patchwork: Building a Big Picture Out of Scraps of Data

Taylor Groves (NERSC – Berkeley, US, tgroves@lbl.gov)

The focus of NERSC is moving from simulations limited to a single site to complex workflows that span multiple sites – incorporating both data center and HPC technologies. This patchwork of telemetry and control creates additional challenges for data analysis. In this lightning talk, I present the perspective from NERSC and highlight challenges and opportunities for future research.

5.2 Challenges for Holistic Monitoring When Attempting Codesign Support

Terry Jones (Oak Ridge National Laboratory – Knoxville, US, trjones@ornl.gov)

For many years, High Performance Computing (HPC) relied on performance improvements based chiefly on frequency scaling; during that era, performance portability was straightforward between different generations of machines. Current computer architecture trends have shifted to other strategies for improvements and exhibit rapidly increasing complexity and extreme heterogeneity to provide performance gains. Codesign activities have emerged as vital processes for mapping applications to the underlying hardware. This talk covers key challenges that arise for holistic monitoring when trying to support codesign.

5.3 Large-scale Vendor-Neutral Power Monitoring

Tapasya Patki (Lawrence Livermore National Laboratory – Livermore, US, patki1@llnl.gov)

Low-level power and performance dials vary significantly from vendor-to-vendor, reducing the portability of system software and monitoring tools. This talk presents Variorum, a library for vendor-neutral power management, and its integration with the Lightweight Distributed Metric System (LDMS).

5.4 Towards an Efficient and Concise Characterization of Temporal I/O Behavior

Francieli Boito (University of Bordeaux/Inria – Bordeux, FR, francieli.zanon-boito@inria.fr)

Existing options for I/O characterization on HPC systems include the creation of complete traces – a log of all I/O operations and their information – or of aggregated statistics, such as the total number and size of accesses. In this talk, I discuss the importance of temporal I/O behavior information and recent advances in using signal processing techniques to obtain it and represent it in a concise manner.

5.5 Containerized Real Time Job Anomaly Detection

Mike Showerman (NCSA – Urbana, US, mung@illinois.edu)

Job anomaly detection allows for identifying irregular job behavior that may affect job execution times, other user's job performance, or system throughput. Examples include load imbalances, stalled jobs, unused requested resources (like GPUs), or high shared resource usage. This talk presents the container-based job anomaly detection that was deployed on the Blue Waters system at NCSA. It allowed for analyzing thousands of jobs in real time on a 27,000 nodes HPC system.

5.6 A Holistic View of Memory Utilization on HPC Systems: Current and Future Trends

Ivy Peng (KTH Royal Institute of Technology – Stockholm, SE, ipeng@acm.org)
Kathleen Shoga (Lawrence Livermore National Laboratory, US, shoga1@llnl.gov)

The Sonar monitoring infrastructure is a central system-monitoring infrastructure deployed at Livermore Computing. Its main components include sample monitoring, data ingestion, staging, archive, and analytics. In this work, we showcase the analysis results from Sonar monitoring infrastructure that can drive the system co-design for the memory subsystem on next generation HPC systems. Memory subsystem is one crucial component of a computing system. Co-designing memory subsystems becomes increasingly challenging as workloads continue evolving on HPC facilities and new architectural options emerge. This work provides a large-scale study of memory utilization with system-level, job-level, temporal and spatial patterns on a CPU-only and a GPU-accelerated leadership supercomputer. From system-level monitoring data that spans three years, we identify a continuous increase in memory intensity in workloads over recent years. We showcase how monitoring data on production systems can reveal different hotspots in memory usage in applications on large-scale systems. We introduce two metrics, "spatial imbalance" and "temporal imbalance", to quantify the imbalanced memory usage across compute nodes and throughout time in jobs. Finally, we identify representative temporal and spatial patterns from real jobs, providing quantitative guidance for research on efficient resource configurations and novel architectural options.

6 Talks: Feedback Driven Response

6.1 Where Rubber Meets the Road: Challenges in Actionable Response to I/O Performance Data

Phil Carns (Argonne National Laboratory – Lemont, US, carns@mcs.anl.gov)

Methods of measuring and recording performance data have been around as long as computers themselves, and the first reaction to any performance data is to imagine how to improve performance. Why isn't this more automated by now? This talk will share my experiences dealing with this problem in the arena of HPC I/O, including challenges in identifying trigger behaviors, selecting responses, and mechanisms for enacting responses. Thoughtful solutions to these problems have the potential for enormous positive impact in scientific computing.

6.2 Driving Response to Uncertainty in Job Duration and Loss of Work

Frédéric Suter (Oak Ridge National Laboratory – Knoxville, US, suterf@ornl.gov)

Uncertainty in job duration, caused by different input parameters or resource sharing can cause a loss of work if the resource reservation is not well dimensioned. To address this issue, several approaches can be followed at different times. Simulation can be used as a comprehensive twin of the job and help to improve the estimation of the job duration before its submission. Uncertainty-aware scheduling strategies can be exploited by the job and resource management system which is the only system component that has a global view of all the jobs and the resources. Finally, by running a job co-pilot that can access both the performance data exported by the application and information coming from the job scheduler, we can better react to unexpected events and prevent unwanted loss of work.

6.3 Hierarchical Control for Runtime Adaptation

Valeria Cardellini (University of Roma Tor Vergata, IT, cardellini@ing.uniroma2.it)

To tackle the fundamental challenges of performance and workload uncertainty, we can exploit a hierarchical control architecture to adapt at runtime the managed system or application. This control pattern avoids the scalability limitations of fully centralized controllers as well as the lack of coordination of fully decentralized schemes. Moreover, it allows for separation of concerns and time scales. As a case study, I present how to manage the horizontal auto-scaling of distributed applications deployed over heterogeneous computing infrastructures by means of heterogeneity-aware adaptation policies that run on a two-layered hierarchy of controllers. Application-level controllers steer the adaptation process for whole applications, aiming to guarantee user-specified Quality of Service requirements while easing the setting of control knobs that are exposed to the users. Lower-layer controllers take auto-scaling decisions for single application components using reinforcement learning techniques. To adopt the latter in an efficient and beneficial way, we can blend together models and learning to improve the learning velocity as well as function approximation techniques to reduce the memory requirements.

6.4 Laying the Foundation for Self-Organizing Systems

Ann Gentile (Sandia National Laboratories – Albuquerque, US, gentile@sandia.gov)

Main reference SAND2023-02343PE

Operational efficiency can be gained through data-driven operations, intelligent resource management, and resiliency handling, however the current data-gathering, communication, and response hooks do not exist! We propose that a fundamental redesign of system software, middleware, applications, and architectures and their interactions is needed to provide the hooks and all the components must be self-aware and self-organizing in order to make the right local and global decisions on the time-scales necessary for effective response.

7 Talks: Analytic and Informatics Approaches

7.1 Time Series Analysis for Performance Monitoring of HPC Systems

Abdullah Mueen (University of New Mexico – Albuquerque, US, mueen@unm.edu)

HPC performance monitoring tasks must respond back to the system and users in order to improve the overall performance, usability, resilience, robustness among many others. Data mining techniques can serve these needs by catering real-time knowledge extraction tools. In this talk, I show streaming data mining tools for pattern matching, clustering and anomaly detection that are data agnostic and applicable to performance monitoring data. I also discuss the potential response mechanisms that can be supported by offline knowledge extraction and online learning techniques with inputs from applications.

7.2 Missing Gaps and Opportunities in HPC Operational Data Analytics

Devesh Tiwari (Northeastern University – Boston, US, d.tiwari@northeastern.edu)

Traditionally, we have assumed that large-scale computing users are fairly boring and that their workloads often do similar things repetitively. But, now things are changing and changing fast. Our workloads and users are becoming interesting and, often, are surprising us with new trends and behavior. In this talk, I will discuss a few lessons I learned as we applied AI/ML methods to HPC resource management problems. In particular, I will discuss performance variability identification, performance auto-tuning, and open-sourcing a large HPC datasource.

7.3 Continuously Detecting Workflow Anomalies using Graph Neural Networks – Lessons Learnt

Krishnan Raghavan (Argonne National Laboratory – Lemont, US, kraghavan@anl.gov)

In this talk, I present the analysis challenges that we face when applying graph neural networks in detecting workflow anomalies. I begin this talk with an overview on graph neural networks and describe the need for them. I then discuss the challenges that are encountered when modeling the data from a workflow as graphs and illustrate the notion of imbalance and noise in this context. I then define how these challenges amplify when more and more data is collected from the system. Moreover, to correct this issue, I illustrate the need for continual learning in this problem and end the talk with the question, "how can we model and correct the challenges due to the drift in the distribution of the data?"

7.4 Best Practices on the Practical Use of Machine Learning in Hybrid Memory Management

Thaleia Doudali (IMDEA – Madrid, ES, thaleia.doudali@imdea.org)

In this talk, I will share some lessons learned on how we can build practical foundations when designing hybrid memory management systems that leverage machine learning methods to improve their effectiveness. Then, I will make a case on how the integration of visualization inside the systems software has potential to accelerate and improve resource management systems. I will end with a crazy idea on how changing data representations into image or text can unlock new opportunities for integrating cross-domain algorithms into systems solutions.

7.5 AI & Analytics in Service of Green Computing

Hilary Egan (NREL – Golden, US, hilary.egan@nrel.gov)

In this talk I present a series of case studies in attempting to use AI/Analytics to improve the sustainability in computing at NREL's ESIF data center, and some future needs and directions for AI algorithms in this space. I begin with an case anomaly detection for advanced cooling systems study done in collaboration with HPE, showing examples on a blower door failure and scale build-up in a heat exchanger. Next I show attempts to optimize system PUE with facility set points and argue we need to integrate digital twin modeling.

I then argue a wide variety of grid-integration technologies will require sophisticated load shaping and present preliminary results in this space. I conclude with discussing future directions in AI including digital twins, modular AI, time-series foundation models.

8 Talks: Actionable Analytics and Response

8.1 AIOPS: Artificial Intelligence for IT Operations in HPC Systems

Jeff Hanson (Hewlett Packard Enterprise – Spring, US, jeff.hanson@hpe.com)
Torsten Wilde (Hewlett Packard Enterprise – Böblingen, DE, torsten.wilde@hpe.com)

Available data for HPC data centers are increasing in velocity and scale. Using traditional threshold based alerting mechanisms or visual inspection of operation graph data are limited in scope and resolution of detection. In this talk we present an implemented framework that uses ML/AI models to work on streaming data to detect anomalous behavior and reduce alarm fatigue for the system administrators. Enabling the research community to develop and implement control loops based on system information is not easy since system integrators provide their own management frameworks to keep the system in a defined and safe operating state. In the second part of this talk we will highlight the challenge using system power and energy management as an example. We will propose different approaches that would enable community developed management and control processes to work in synergy with system integrator provided mechanisms.

8.2 Energy Management on HPC Systems

Oriol Vidal (Barcelona Supercomputing Center, ES, oriol.vidal@bsc.es)

Full power management of large infrastructures requires an autonomous way to monitor system running jobs, as well as identifying those power saving opportunities. This talk will present how the EAR software fills ODA framework's gaps concerning HPC power and energy management, from basic monitoring and accounting to power saving strategies during application runtime and cluster-level power-cap.

9 Talks: Autonomous Loop Lightning Talks

9.1 Malleability, Monitoring and Modeling

Isaías A. Comprés Ureña (TU München – Garching, DE, compresu@in.tum.de)

The priority-based First-Come First-Serve (FCFS) with backfilling heuristic has been very successful in maximizing node utilization. However, since node utilization depends on system state and user provided jobs that are queued, it is not always possible to reach close to full node utilization. We propose malleability as a means to add additional opportunities for schedulers to reach near full node utilization. With malleability, nodes can be added or removed from jobs. Monitoring and modeling techniques are needed, to ensure quality decisions are made.

9.2 Control Loops in the PaPP Project

Sven Karlsson (Technical University of Denmark – Kongens Lyngby, DK, svea@dtu.dk)

Since the improvement of computer performance can no longer be satisfied by simply raising clock frequencies, architectures are evolving towards both the multiplication of processing elements and heterogeneity of their functions. The aforementioned execution platforms combined with steadily exacting non-functional performance requirements such as execution speed, timeliness, and power consumption challenges traditional design methods. The PaPP project included work on many aspects including improved resource management, adaptivity in all parts of the technology stack including the applications, and a control loop for managing non-functional performance requirements. The PaPP software infrastructure and the control loops were briefly introduced with demonstrations presented.

9.3 Job Efficiency Reporting

Thomas Jakobsche (University of Basel, CH, thomas.jakobsche@unibas.ch)

HPC systems provide high computing power to different domain scientists. Not all users possess the necessary knowledge and experience to fully utilize the capabilities of HPC systems. Certain users tend to submit inaccurate resource requests, leading to inefficient job scheduling, longer job wait times, and ultimately to wasted computing resources. HPC systems offer several tools to support users in order to improve their resource requests (e.g. SLURM's `seff` and `sacct` commands). However, these commands are missing easily digestible job statistics that can provide, for example, the longest execution time for a group of jobs. This lightning talk presented a command-line functionality based on the SLURM `sacct` command that introduces grouping and tabular-like functionality to job accounting data, that is easily available to administrators and users. This tool can support administrators and users to quickly identify groups of jobs with inaccurate resource requests, create job efficiency reports to raise awareness, and support users to improve future resource requests based on previous job submissions.

9.4 Code Path and Parameter Selection for Compute Loops

Thomas Gruber (Universität Erlangen-Nürnberg, DE, thomas.gruber@fau.de)

Applications for ODEs and PDEs often benefit from different coder variants but performance is highly dependent on the input. For testing these variants as well as hardware and runtime parameters different combinations need to be tested. By letting a genetic algorithm select testable parameter combinations, the amount of combinations can be reduced to a reasonable set. With measurements, the combinations can be graded and the best one selected for the given input. With refinements from run to run, the application can be executed with little disturbance and still find optimal execution settings. The user interacts with the library by supplying code variants through code instrumentation.

9.5 Feedback Loop for Mitigating Assignment of Slow OSTs for File I/O

Jim Brandt (Sandia National Laboratories – Livermore, and Albuquerque, US, brandt@sandia.gov)

This presentation was about an explicit use case at Sandia National Laboratories where processes of an application can be handed Lustre Object Store Targets (OSTs) with significantly poorer performance than other processes. The worst performance gates the application performance and in the presented case significantly increases the application run time (e.g., >50%). The mitigating approach being explored is creation of a feedback loop from the monitoring and analysis system to the application which could then take the action of deleting and re-creating the file it is writing to and ideally be provided with more performant OSTs.

10 Fundamental Results

Bringing together experts from different disciplines was needed to address the wide-ranging and interdependent aspects of the problem space of HPC Operations and enabled us to create a community and network around moving towards autonomous HPC system management and operation.

We discussed the fundamental requirements and gaps to enable autonomous feedback and response loops in production HPC systems. These are summarized, at a high level, below.

10.1 Assessment of the state of autonomous feedback and response loops in HPC operations

The community view was that progress on autonomous feedback and response loops has been limited. Individual participants have implemented feedback and response loops (either with or without human-in-the-loop) in order to improve energy efficiency; to improve performance of applications in the face of shared resources, such as networks and file systems; to diagnose performance and configuration issues, and to optimize competing goals of cost vs. performance. Many of these efforts addressed specific cases, specific architectures, and/or were designed to demonstrate potentials. While many of these have been successful in terms of their limited goals, they have not typically resulted in continuous production deployment to affect or improve operations. We believe that more *holistic* design and approaches to feedback and response are needed due to the interdependencies of systems, subsystems, and applications in order to fully benefit from autonomous feedback in HPC operations.

10.2 Assessment of most important challenges and opportunities

The community identified a number of major *challenges* that have impeded progress in the development of autonomous feedback and response loops in HPC operations. The HPC systems domain innately consists of complex architectures, the monitoring of which produces high-dimensional time-series data, making actionable analytics difficult. Effective feedback and response hooks are limited or require privileged access. It is difficult to build the required verification and trust necessary to deploy prototypes that alter yet improve system and application behavior on production systems. Without a clear path forward for development and deployment, many efforts have been independent, resulting in implementations that are not interoperable, are not interchangeable, are limited in scope, and are difficult to maintain.

There exists the *opportunity* to leverage the community's experience in feedback and response loops to develop infrastructure and standards for continuous loops to be deployed in production HPC systems that can both facilitate community development and affect production systems.

10.2.1 Conventions to Facilitate Autonomous Feedback and Response Loop Development

We identified the *opportunity* to leverage existing formalism, established in the fields of Self-Adaptive Systems and Autonomic Computing, i.e., the *Monitor, Analyze, Plan and Execute (MAPE), with shared Knowledge (MAPE-K)* control loop [Kephart and Chess, The Vision of Autonomic Computing, 2003], because of the similarity of concepts and interplay of *monitoring, analysis, feedback, and response* in our domain. Arbitrarily complex autonomous actions can be supported by different decentralized architectural design patterns comprising the MAPE-K components. A number of such patterns and heuristics for their application have been established. By leveraging the MAPE-K formalism, we hope to take advantage of the established design patterns and thereby facilitate the development of autonomous loops and the required loop infrastructure that would enable holistic feedback and response for HPC operations.

We discussed that several MAPE-K autonomy control loops can be defined in HPC systems, for example one in each subsystem. Each autonomy loop can be implemented as a centralized, hierarchical, or distributed loop, and can connect to other autonomy loops in the system.

Additionally, we identified the *challenge* of standardizing the interfaces between the different major control loop components, which would allow for easy interchange of components and for collaborative development. The community identified a potential *opportunity* for the development of conventions. This will be realized in a community technical path forward (Section 11.1) where development of the loop components, which include supporting analysis methodologies and decision-making logic, will be used to extract and define system software (e.g., APIs, and modular design) and data engineering (e.g., data formats and data requirements) conventions that will be adopted by the community.

10.2.2 Open data sets/challenges

An *opportunity* exists to build and leverage the external community via the application and development of analysis for our domain. The HPC systems community has substantial production telemetry datasets. While there exists some understanding of our architectures and applications, there are still substantial limitations in the description of our data and our understanding of how that data reflects performance issues and headroom opportunities. This has complicated the application of analytics in our domain. Statistical and Machine Learning algorithms may be suited to overcome these *challenges*. However, the lack of expertise in these techniques in the systems community has limited our ability to make analytic progress on our own. By building a community that includes not only HPC experts but also analysis researchers, further progress may be made.

Such collaboration and progress will be non-trivial. The processes of producing data and analyzing data are not independent. Significant exchange of knowledge will be required to ensure uniform understanding of data and expectations of metrics and features of importance. To be of use, telemetry data should include semantic labels so that not just names, but also understanding of meaning, is conveyed. Datasets need to include events of interest and with frequencies and details that meet the requirements of algorithmic application (e.g., evidence of sufficient rare events, unanticipated bias features in the data). Additionally, a domain expert's identification of expected features of significance and what they functionally relate to in the operation of an HPC system would be of immense benefit to consumers of this data. System state (idle vs. busy with well-known benchmarks/proxies/canaries running) and changes in state should be provided over the time window of the supplied data set.

The participants explored that an *opportunity* to facilitate this could be through the development of an open dataset and challenge where we would test the readiness of and refine the requirements for the monitoring data produced and collected on the systems to be suitable for analysis by AI/ML/statistical methods.

The participants also agreed that the community would benefit from a survey of existing data analysis methods and data sets with the properties described above.

11 Planned Paths Forward

At the seminar, we laid the groundwork for continuing progress as a community on three fronts: technical, interaction, and funding. These are discussed in this section.

11.1 Identification of a community collaborative technical path forward

To drive the community along a collaborative technical path forward, we discussed various use cases that are commonly encountered across HPC systems and sites. Three use cases were identified as representative of burning problems for which the community could develop proof-of-concept solutions. The development process would enable us to explore options for architectural design patterns and interfaces and their extensibility to support multiple cases. Successful development could then drive technical change in both loop development and production adoption, form a foundation for a community vision report, and serve as the basis for future funding proposals.

These use cases are:

1. Interaction of an HPC monitoring and analysis system with the scheduler on behalf of running applications in order to ensure that a reasonably progressing application is not terminated prematurely. The autonomous loop would involve continuous monitoring and analysis to project an application's execution time, which may vary due to system conditions, competing workload, or problem specifics, and feedback to extend the allowed running time. This use case is representative of instances where feedback would need to be directed to system software and middleware.
2. Interaction of an HPC monitoring and analysis system with applications in a feedback loop to inform a target application of a problem for which it could take mitigating action. This use case is related to the above case, but is representative of instances where feedback is directed to applications.
3. Detecting performance degrading misconfiguration of an application deployed on HPC resources. The autonomous loop would involve performance assessment and attribution of existing configurations to provide feedback to humans and applications of potential misconfigurations. Distinctions in this use case, as opposed to the above two, are the source of the condition of interest and the feedback direction and target.

Participants have agreed to further define the scope and parameters of these use cases and devote resources to collaboratively develop proof-of-concept-grade solutions. These solutions will inform further understanding of how to generalize approaches to solving the classes of problems represented and to extract commonalities that we will use to define conventions for development. Sites with HPC resources that would be open to deploying exploratory autonomous feedback and response loops for controlling operations and that could potentially host external developers were identified. These include potential large-scale, end-of-life, systems. These sites will be looking further into details for the necessary arrangements.

Note that changing the HPC operational model to incorporate pervasive autonomy loops is not merely a software development and data engineering challenge. Significant work will remain to explore issues that need to be overcome for system administrator and site acceptance and adoption. These include identifying and granting authority for control, establishing scope of decisions made, avoiding competing and thrashing responses, and quantifying impact and ensuring response on meaningful timescales.

11.2 Identification of community interaction path forward

Several opportunities to continue interaction and collaboration within the community and make technical progress on improving the performance of HPC applications and efficiency and sustainability of HPC systems as a community have been identified and agreed upon:

- Setting up a regular virtual meeting to continue the discussions started at Dagstuhl. Progress on the proof-of-concept autonomous control will also be coordinated at these meetings.
- Further meetings at other community-relevant events (ISC/MODA, SC/Quantitative Supercomputing Codesign, IEEE Cluster/HPCMASPA conferences, and EEHPCWG Workshop) are envisioned to drive further discussions, finalize, and disseminate the community vision report.
- A proposal for a Birds-of-a-Feather (BoF) session at SC'23 and/or a position paper to a workshop or symposium will be submitted to widen community awareness.
- The organization of a Special Issue at a leading journal will also be explored as a way to expand technical and community knowledge of the field.

11.3 Identification of community collaborative funding path forward

The community is spread all over the world, with most participants of the seminar being either US- or Europe-based. As there is interest to collaborate further on the topics discussed during the seminar, many attendees showed interest in submitting a joint proposal to fund further activities to drive the development of autonomous feedback and response loops for HPC operations forward. However, as a joint EU/US funding call that allows for project partners from both continents seems out of reach, the attendees agreed to continue with separate proposals in the US and EU that would support each other and include close collaboration between the two. Additionally, other funding instruments that would allow for bilateral collaboration have been identified.

12 Acknowledgments

The organizers would thank Schloss Dagstuhl and their staff for hosting and making this seminar possible.

Some of the contributors to this research were supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration.

A portion of this work was supported by the Scientific Discovery through Advanced Computing (SciDAC) program funded by the U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research (ASCR) under contract DE-SC0021299.

Some of the participants received funding from the "Adaptive multitier intelligent data manager for Exascale (ADMIRE)" project, funded by the European Union's Horizon 2020 JTI-EuroHPC Research and Innovation Programme (grant 956748).

Some of the participants received funding from the "DAPHNE: Integrated Data Analysis Pipelines for Large-Scale Data Management, HPC, and Machine Learning" project, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 957407.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

Participants

- Francieli Boito
INRIA – Bordeaux, FR
- Jim Brandt
Sandia National Labs – Albuquerque, US
- Valeria Cardellini
University of Rome "Tor Vergata", IT
- Philip Carns
Argonne National Laboratory, US
- Florina M. Ciorba
Universität Basel, CH
- Isaías Alberto Comprés Ureña
TU München – Garching, DE
- Thaleia Dimitra Doudali
IMDEA Software Institute – Madrid, ES
- Hilary Egan
NREL – Golden, US
- Ahmed Eleliemy
Universität Basel, CH
- Ann Gentile
Sandia National Labs – Albuquerque, US
- Taylor Groves
Lawrence Berkeley National Laboratory, US
- Thomas Gruber
Universität Erlangen-Nürnberg, DE
- Jeff Hanson
HPE – Lakewood, US
- Utz-Uwe Haus
HPE HPC/AI EMEA Research Lab – Wallisellen, CH
- Esa Heiskanen
CSC Ltd. – Kajaani, FI
- Kevin A Huck
University of Oregon – Eugene, US
- Thomas Ilsche
TU Dresden, DE
- Thomas Jakobsche
Universität Basel, CH
- Terry Jones
Oak Ridge National Laboratory, US
- Sven Karlsson
Technical University of Denmark – Lyngby, DK
- Allen D. Malony
University of Oregon – Eugene, US
- Henrique Mendonça
CSCS – Lugano, CH
- Abdullah Mueen
University of New Mexico, US
- Michael Ott
LRZ – München, DE
- Tapasya Patki
LLNL – Livermore, US
- Ivy Bo Peng
KTH Royal Institute of Technology – Stockholm, SE
- Krishnan Raghavan
Argonne National Laboratory, US
- David Schibeci
Pawsey Supercomputing Centre – Kensington, AU
- Kathleen Shoga
LLNL – Livermore, US
- Michael Showerman
University of Illinois at Urbana-Champaign, US
- Frédéric Suter
Oak Ridge National Laboratory, US
- Oriol Vidal
Barcelona Supercomputing Center, ES
- Torsten Wilde
HPE- Böblingen, DE
- Keiji Yamamoto
RIKEN – Hyogo, JP

Remote Participants

- Devesh Tiwari
Northeastern University –
Boston, US