Report from Dagstuhl Seminar 23111

Computational Complexity of Discrete Problems

Anna Gál*[1], Meena Mahajan*[2], Rahul Santhanam*[3], Till Tantau*[4], and Manaswi Paraashar†[5]

1 **University of Texas – Austin, US.** panni@cs.utexas.edu
2 **The Institute of Mathematical Sciences – Chennai, IN.** meena@imsc.res.in
3 **University of Oxford, GB.** rahul.santhanam@cs.ox.ac.uk
4 **Universität zu Lübeck, DE.** tantau@tcs.uni-luebeck.de
5 **Aarhus University, DK.** manaswi.isi@gmail.com

Abstract

This report documents the program and activities of Dagstuhl Seminar 23111 “Computational Complexity of Discrete Problems”, which was held in-person in March 2023 (the previous instance of the seminar series had been held online in March 2021). Following a description of the seminar’s objectives and its overall organization, this report lists the different major talks given during the seminar in alphabetical order of speakers, followed by the abstracts of the talks, including the main references and relevant sources where applicable. The return to an in-person setting allowed an intense atmosphere of active research and interaction throughout the five day seminar.

Seminar March 12–17, 2023 – https://www.dagstuhl.de/23111
2012 ACM Subject Classification Theory of computation → Probabilistic computation; Theory of computation → Streaming models; Theory of computation → Circuit complexity; Theory of computation → Pseudorandomness and derandomization
Keywords and phrases circuit complexity, communication complexity, computational complexity, lower bounds, randomness
Digital Object Identifier 10.4230/DagRep.13.3.17

1 Executive Summary

Anna Gál (University of Texas, Austin, US)
Meena Mahajan (The Institute of Mathematical Sciences, Chennai, IN)
Rahul Santhanam (University of Oxford, GB)
Till Tantau (Universität zu Lübeck, DE)

Computational complexity studies the amount of resources (such as time, space, randomness, parallelism, or communication) that are necessary to solve computational problems in various models of computation. Finding efficient algorithms for solving computational tasks is crucial in many practical applications. Despite a long line of research, for many discrete problems that arise in practice it is not known if they can be solved efficiently – in particular, in (randomized) polynomial time. While efficient algorithms clearly have obvious applications, knowing that a problem can*not* be solved efficiently can *also* have high practical impact. For example, lower bounds on the amount of resources needed to solve specific problems can be used to construct good pseudorandom generators to derandomize probabilistic algorithms. Similarly, the security of our currently used crypto-systems hinges on the *assumption* that

* Editor / Organizer
† Editorial Assistant / Collector

certain discrete problems – like factoring – are *hard* to solve; and we would very much like to *prove* that more efficient algorithms cannot exist for factoring. Proving lower bounds is a challenging task since one needs to argue against all possible algorithms. In the last few decades, lower bound methods have been developed for various restricted or special models of computation. These results often involve the use of sophisticated mathematical techniques and despite a lot of effort we still do not have – somewhat frustratingly – strong enough techniques to establish for instance superlinear lower bounds for specific problems in general computational models, such as the model of Boolean circuits. In this Dagstuhl Seminar, which is the most recent incarnation of a seminar series that stretches back many years, we brought together leading experts and talented junior researchers to discuss the most exciting recent developments in different areas of computational complexity of discrete problems – both regarding recent *results,* but also regarding *open problems.* In both cases, a particular focus was on lower bounds and on whether and, if so, how ideas and methods from one theory area can yield insights in another theory area.

To enable and encourage discussions between the researchers present in Dagstuhl, time was allotted to three different formats: The presentation and discussion of current research results and methods, the presentation and discussion of open problems and conjectures, and on-site collaborative theory research. Each day of the seminar started with a morning session dedicated to survey talks, talks sharing research results, and talks introducing new techniques (the titles and abstracts of most of these talks appear later in this report). The afternoons were dedicated to collaborative research in various forms: On Tuesday, research was done in break-out sessions in which smaller groups of participants explored different open research questions in depth. The topics covered were the following (in alphabetical order of chairs):

1. *Tree Codes,* chaired by Gil Cohen.
2. *Lifting Dichotomies,* chaired by Yuval Filmus.
3. *Finding Tarski Fixed-Points,* chaired by Kristoffer Hansen.
4. *Hitting Sets Versus Orthogonal Vectors (a.k.a. kSAT Versus $\forall\exists k$SAT),* chaired by Marvin Kühnemann.
5. *Simple Versions of #SAT,* chaired by Till Tantau.

Two other formats were intended to intrigue participants in research questions beyond their own speciality (and succeeded in doing so). Firstly, there were open problem sessions, where Gil Cohen, Yuval Filmus, Mika Göös, Rohit Gurjar, Alexander Kulikov, Jakob Nordström, Rüdiger Reischuk, Robert Robere, and Ben Lee Volk presented different research questions. Secondly, at various points during the seminar, there were short "talks to talk about", which introduced exciting recent topics, results, or problems and which gave people intriguing topics to discuss over lunch or dinner. Talks to talk about were given by Sourav Chakraborty, Manaswi Parashar, and Till Tantau.

A final important objective of the seminar was to foster collaborations not only between researchers working on different topics, but also between junior and senior researchers. Towards this aim, talks of more junior researchers were scheduled (as far as possible) on the first day, giving them early exposure and allowing other participants to talk to them about the presented research results during the whole week. Naturally, both junior and senior participants had ample opportunity to socialize, be it during the traditional Wednesday afternoon hike or the wine-and-cheese party on Thursday.

The organizers, Anna Gál, Meena Mahajan, Rahul Santhanam, and Till Tantau, thank all participants for the many contributions they made. We would also like to especially thank the Dagstuhl staff, who were – as usual – extremely friendly, helpful, and professional regarding all organizational matters surrounding the seminar. Finally, we express our great gratitude to Manaswi Paraashar for his invaluable help assembling and preparing this report.

2 Table of Contents

Executive Summary

Overview of Talks

Working groups

Open problems

3 Overview of Talks

3.1 Models of CDCL solving for quantified Boolean formulas

Olaf Beyersdorff (Friedrich-Schiller-Universität Jena, DE)

Joint work of Benjamin Böhm, Tomas Peitl, Olaf Beyersdorff
Main reference Benjamin Böhm, Tomás Peitl, Olaf Beyersdorff: "QCDCL with Cube Learning or Pure Literal Elimination – What is Best?", in Proc. of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022, pp. 1781–1787, ijcai.org, 2022.
URL https://doi.org//10.24963/ijcai.2022/248

This talk explained the relations between solvers based on the conflict-driven clause learning (CDCL) paradigm for quantified Boolean formulas (QBF) and QBF resolution systems. Particular emphasis was placed on how to theoretically model CDCL algorithms for QBF and investigate the proof-theoretic strength of different QCDCL solving approaches.

3.2 Distinct Elements in Streams: An Algorithm for the Text Book

Sourav Chakraborty (Indian Statistical Institute – Kolkata, IN)

Joint work of Sourav Chakraborty, Kuldeep Meel, N.V. Vinodchandran
Main reference Sourav Chakraborty, N. V. Vinodchandran, Kuldeep S. Meel: "Distinct Elements in Streams: An Algorithm for the (Text) Book", CoRR, Vol. abs/2301.10191, 2023.
URL https://doi.org//10.48550/arXiv.2301.10191

Given a data stream $D = a_1, a_2, ..., a_m$ of m elements where each $a_i \in [n]$, the Distinct Elements problem is to estimate the number of distinct elements that appear in the stream.

Distinct Elements has been a subject of theoretical and empirical investigations over the past four decades resulting in space optimal algorithms for it. All the current state-of-the-art algorithms are, however, beyond the reach of an undergraduate textbook owing to their reliance on the usage of notions such as pairwise independence and universal hash functions. We present a simple, intuitive, sampling-based space-efficient algorithm whose description and the proof are accessible to undergraduates with the knowledge of basic probability theory.

3.3 Testing correctness of samplers using property testing: from theory to practice and back again

Sourav Chakraborty (Indian Statistical Institute – Kolkata, IN)

Joint work of Sourav Chakraborty, Kuldeep Meel, Yash Pote, Rishiraj Bhattacharyya, Priyanka Golia, Mate Soos

How can one test the correctness of a program that is supposed to output an element from a large universe according to a certain distribution? These kind of programs are heavily used in real life but are rarely tested for correctness.

This problem can be framed as a problem in property testing. Property testing is a subject that deals with these challenges. It tries to design sub-linear algorithms for testing various properties of inputs. The key lies in the way the data is accessed by the algorithm.

One of the central problems in property testing and many other related subjects is testing if a distribution has a certain property – say whether a distribution on a finite set is uniform. The conventional way of accessing the distributions is by drawing samples according to the distributions. Unfortunately, in this setting the number of samples that are necessary for testing properties of distribution (for most natural properties) is polynomial in the size of support of the distribution. Thus when the support is relatively big the algorithms become impractical in real life applications.

We define a new way of accessing the distribution using "conditional-sampling oracle". This oracle can be used to design much faster algorithms for testing properties of distribution and thus makes the algorithm useful in practical scenarios.

We show that the conditional oracle can be implemented in many real life problems and we have been able to show the usefulness of this model and our algorithms in practical purposes and in other areas of research – like testing of probabilistic verification. This model also throws a number of interesting theoretical questions.

The talk will be based on the following works:

References

1 Eldar Fischer, Arie MAtsliah and Yonatan Goldhrish: On the Power of Conditional Samples in Distribution Testing, (SICOMP 2016)
2 Rishiraj Bhattacharyya: Property Testing of Joint Distributions using Conditional Samples, (ToCT 2018)
3 Kuldeep Meel: On Testing of Uniform Samplers, (AAAI2019)
4 Kuldeep Meel and Yash Pote: On Testing of Samplers, (NeuRIPS 2020)
5 Kuldeep Meel, Priyanka Golia and Mate Soos: Designing Samplers is Easy: The Boon of Testers, (FMCAD22)
6 Kuldeep Meel, Priyanka Golia and Mate Soos: On Quantitative Testing of Samplers, (CP22)
7 Ansuman Banerjee, Shayak Chakraborty, Sayantan Sen, Uddalok Sarkar and Kuldeep Meel: Testing of Horn Samplers, (AISTAT 2023)

3.4 Graph Colouring Is Hard on Average for Polynomial Calculus

Susanna de Rezende (Lund University, SE), Jakob Nordström (University of Copenhagen, DK & Lund University, SE)

Joint work of Jonas Conneryd, Susanna de Rezende, Shuo Pang, Jakob Nordström, Kilian Risse

We prove that polynomial calculus and hence also Nullstellensatz requires linear degree to refute that sparse random regular graphs, as well as sparse Erdös-Rényi random graphs, are 3-colourable. Using the known relation between size and degree for polynomial calculus proofs, this implies strongly exponential lower bounds on proof size.

3.5 The HITTING proof system

Yuval Filmus (Technion – Haifa, IL)

Joint work of Yuval Filmus, Edward A. Hirsch, Artur Riazanov, Alexander Smal, Marc Vinyals
Main reference Yuval Filmus, Edward A. Hirsch, Artur Riazanov, Alexander Smal, Marc Vinyals: "Proving Unsatisfiability with Hitting Formulas", CoRR, Vol. abs/2302.06241, 2023.
URL https://doi.org//10.48550/arXiv.2302.06241

A tree-like Resolution refutation of a CNF is a decision tree that solves the falsified clause search problem. We can think of the leaves of the decision tree as a partition of the space of truth assignments into "monochromatic" subcubes, in the sense that each subcube can be associated with a single refuted clause. A HITTING refutation of a CNF is any partition of the space of truth assignments into monochromatic subcubes.

We explore the relation between HITTING and other proof systems. By construction, HITTING p-simulates tree-like Resolution, and in contrast, tree-like Resolution qp-simulates HITTING, and there is a qp-separation between the two systems. Resolution can be exponentially more powerful than HITTING, but we conjecture that it does not p-simulate HITTING. Using the Raz-Shpilka PIT, we show that Extended Resolution p-simulates HITTING, though this is probably an overkill.

3.6 Top-Down Lower Bounds for Depth-Four Circuits

Mika Göös (EPFL Lausanne, CH)

Joint work of Mika Göös, Artur Riazanov, Anastasia Sofronova, Dmitry Sokolov

We present a top-down lower-bound method for depth-4 boolean circuits. In particular, we give a new proof of the well-known result that the parity function requires depth-4 circuits of size exponential in $n^{1/3}$. Our proof is an application of robust sunflowers and block unpredictability.

3.7 Capturing one-way functions via NP-hardness of meta-complexity

Shuichi Hirahara (National Institute of Informatics – Tokyo, JP)

Main reference Shuichi Hirahara: "Capturing One-Way Functions via NP-Hardness of Meta-Complexity", in Proc. of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, p. 1027–1038, Association for Computing Machinery, 2023.
URL https://doi.org//10.1145/3564246.3585130

We present the first characterization of a one-way function by worst-case hardness assumptions: A one-way function exists iff NP is hard in the worst case and "distributional Kolmogorov complexity" is NP-hard under randomized reductions. Here, the t-time bounded distributional Kolmogorov complexity of a string x given a distribution D is defined to be the length of a shortest t-time program that outputs x given as input y drawn from the distribution D with high probability. The characterization suggests that the recent approaches of using meta-complexity to exclude Heuristica and Pessiland are both sufficient and necessary.

3.8 Unprovability of strong complexity lower bounds in bounded arithmetic

Igor Carboni Oliveira (University of Warwick – Coventry, GB)

Joint work of Igor Carboni Oliveira, Jiatu Li
Main reference Jiatu Li, Igor Carboni Oliveira: "Unprovability of strong complexity lower bounds in bounded arithmetic", Electron. Colloquium Comput. Complex., Vol. TR23-022, 2023.
URL https://eccc.weizmann.ac.il/report/2023/022

While there has been progress in establishing the unprovability of complexity statements in lower fragments of bounded arithmetic, understanding the limits of Jeřábek's theory APC_1 (2007) and of higher levels of Buss's hierarchy S^i_2 (1986) has been a more elusive task. Even in the more restricted setting of Cook's theory PV_1 (1975), known results often rely on a less natural formalization that encodes a complexity statement using a collection of sentences instead of a single sentence. This is done to reduce the quantifier complexity of the resulting sentences so that standard witnessing results can be invoked.

In this work, we establish unprovability results for stronger theories and for sentences of higher quantifier complexity. In particular, we unconditionally show that APC_1 cannot prove strong complexity lower bounds separating the third level of the polynomial hierarchy. In more detail, the lower bound sentence refers to the non-uniform setting ($\exists\forall\exists$ Circuits vs. $\forall\exists\forall$ Circuits) and to a mild average-case lower bound for polynomial size circuits against sub-exponential size circuits.

Our argument employs a convenient game-theoretic witnessing result that can be applied to sentences of arbitrary quantifier complexity. We combine it with extensions of a technique introduced by Krajíček (2011) that was recently employed by Pich and Santhanam (2021) to establish the unprovability of lower bounds in PV_1 and in a fragment of APC_1.

3.9 On small-depth Frege proofs for PHP

Johan Håstad (KTH Royal Institute of Technology – Stockholm, SE)

We study Frege proofs for the one-to-one graph Pigeon Hole Principle defined on the $n \times n$ grid where n is odd. We are interested in the case where each formula in the proof is a depth d formula in the basis given by $\wedge$, $\vee$, and $\neg$. We prove that in this situation the proof needs to be of size exponential in $n^{\Omega(1/d)}$. If we restrict the size of each line in the proof to be of size M then the number of lines needed is exponential in $n/(\log M)^{O(d)}$. The main technical component of the proofs is to design a new family of random restrictions and to prove the appropriate switching lemmas.

3.10 The Elliptic Curve Fast Fourier Transform (ECFFT)

Swastik Kopparty (University of Toronto, CA)

Joint work of Eli Ben-Sasson, Dan Carmon, Swastik Kopparty, David Levit
Main reference Eli Ben-Sasson, Dan Carmon, Swastik Kopparty, David Levit: "Elliptic Curve Fast Fourier Transform (ECFFT) Part I: Low-degree Extension in Time $O(n \log n)$ over all Finite Fields", in Proc. of the 2023 ACM-SIAM Symposium on Discrete Algorithms, SODA 2023, Florence, Italy, January 22-25, 2023, pp. 700–737, SIAM, 2023.
URL https://doi.org//10.1137/1.9781611977554.ch30
Main reference Eli Ben-Sasson, Dan Carmon, Swastik Kopparty, David Levit: "Scalable and Transparent Proofs over All Large Fields, via Elliptic Curves – (ECFFT Part II)", in Proc. of the Theory of Cryptography – 20th International Conference, TCC 2022, Chicago, IL, USA, November 7-10, 2022, Proceedings, Part I, Lecture Notes in Computer Science, Vol. 13747, pp. 467–496, Springer, 2022.
URL https://doi.org//10.1007/978-3-031-22318-1_17

This is based on the papers ECFFT I (Fast algorithms for polynomials over all fields) and ECFFT II (Scalable and Transparent proofs over all large fields), both joint work with Eli Ben-Sasson, Dan Carmon, and David Levit

I will talk about a variant (the ECFFT) of the FFT which is based on elliptic-curve groups in place of multiplicative groups. While the classical FFT over finite fields is directly applicable only when the size of the multiplicative group of the field is special, the ECFFT turns out to be directly applicable over all finite fields (because all finite fields have *some* elliptic curve group whose size is special).

We then use the ECFFT in place of the FFT for applications in fast polynomial algorithms and interactive property testing.

3.11 Locally consistent decomposition of strings with applications to edit distance sketching

Michal Koucký (Charles University – Prague, CZ)

Joint work of Sudatta Bhattacharya, Michal Koucký
Main reference Sudatta Bhattacharya, Michal Koucký: "Streaming k-edit approximate pattern matching via string decomposition", CoRR, Vol. abs/2305.00615, 2023.
URL https://doi.org//10.48550/arXiv.2305.00615

We present a new locally consistent decomposition of strings. Each string x is decomposed into blocks that can be described by grammars of size $\widetilde{O}(k)$ (using some amount of randomness). If we take two strings x and y of edit distance at most k then their block decomposition uses the same number of grammars and the i-th grammar of x is the same as the i-th grammar of y except for at most k indexes i. The edit distance of x and y equals to the sum of edit distances of pairs of blocks where x and y differ. Our decomposition can be used to design a sketch of size $\widetilde{O}(k^2)$ for edit distance, and also a rolling sketch for edit distance of size $\widetilde{O}(k^2)$. The rolling sketch allows to update the sketched string by appending a symbol or removing a symbol from the beginning of the string.

3.12 Polynomial formulations as a barrier for reduction-based hardness proofs

Alexander S. Kulikov (JetBrains Research – Paphos, CY)

Joint work of Tatiana Belova, Alexander Golovnev, Alexander S. Kulikov, Ivan Mihajlin, Denil Sharipov
Main reference Tatiana Belova, Alexander Golovnev, Alexander S. Kulikov, Ivan Mihajlin, Denil Sharipov: "Polynomial formulations as a barrier for reduction-based hardness proofs", in Proc. of the 2023 ACM-SIAM Symposium on Discrete Algorithms, SODA 2023, Florence, Italy, January 22-25, 2023, pp. 3245–3281, SIAM, 2023.
URL https://doi.org//10.1137/1.9781611977554.ch124

The Strong Exponential Time Hypothesis (SETH) asserts that for every $\varepsilon > 0$ there exists k such that k-SAT requires time $(2 - \varepsilon)^n$. The field of fine-grained complexity has leveraged SETH to prove quite tight conditional lower bounds for dozens of problems in various domains and complexity classes, including Edit Distance, Graph Diameter, Hitting Set, Independent Set, and Orthogonal Vectors. Yet, it has been repeatedly asked in the literature whether SETH-hardness results can be proven for other fundamental problems such as Hamiltonian Path, Independent Set, Chromatic Number, MAX-k-SAT, and Set Cover.

In this paper, we show that fine-grained reductions implying even λ^n-hardness of these problems from SETH for *any* $\lambda > 1$, would imply new circuit lower bounds: super-linear lower bounds for Boolean series-parallel circuits or polynomial lower bounds for arithmetic circuits (each of which is a four-decade open question).

We also extend this barrier result to the class of parameterized problems. Namely, for every $\lambda > 1$, we conditionally rule out fine-grained reductions implying SETH-based lower bounds of λ^k for a number of problems parameterized by the solution size k.

Our main technical tool is a new concept called polynomial formulations. In particular, we show that many problems can be represented by relatively succinct low-degree polynomials, and that any problem with such a representation cannot be proven SETH-hard (without proving new circuit lower bounds).

3.13 Colourful TFNP and Propositional Proofs

Robert Robere (McGill University – Montréal, CA)

Joint work of Ben Davis (McGill University), Robert Robere (McGill University)
Main reference Ben Davis and Robert Robere. "Colourful TFNP and Propositional Proofs". To appear at CCC 2023.
URL https://eccc.weizmann.ac.il/report/2023/068/

Recent work in proof complexity has shown that studying many of the major proof systems studied in practice is, in a sense, completely equivalent to studying black-box versions of syntactically-defined subclasses of TFNP. Many weak proof systems, such as Resolution, Sherali-Adams, and Nullstellensatz are now known to admit characterizations of this type, and these new characterizations have been used to obtain new results in both proof complexity and the study of TFNP.

In this talk, we outline recent work in which we have characterized stronger proof systems – including *Res*(k) and higher-depth analogues of Sherali-Adams – inside of TFNP by using the so-called "coloured" generalization of standard TFNP classes. This talk is based on joint work with Ben Davis.

3.14 Simple, deterministic, and fast (but weak) approximation for Edit Distance and Dyck Edit Distance

Michael E. Saks (Rutgers University – Piscataway, US)

Joint work of Michal Koucký, Michael E. Saks
Main reference Michal Koucký, Michael E. Saks: "Simple, deterministic, fast (but weak) approximations to edit distance and Dyck edit distance", in Proc. of the 2023 ACM-SIAM Symposium on Discrete Algorithms, SODA 2023, Florence, Italy, January 22-25, 2023, pp. 5203–5219, SIAM, 2023.
URL https://doi.org//10.1137/1.9781611977554.ch188

The edit distance between two strings, equal to the minimum number of operations (insertions, deletions or substitutions) needed to transform one to the other, is a standard measure of similarity of strings. The classic dynamic programming algorithm for edit distance requires time quadratic in n to compute the edit distance between two strings of length n, and there is evidence (via the strong exponential time hypothesis) that it may be impossible to improve substantially on this time complexity.

Recently, there has been considerable progress in developing approximation algorithms for edit distance (and the more general problem of Dyck edit distance) that are fast and have constant, or near constant approximation factors. These algorithms run in near linear time, but are logically complex, and the constants in both the running time and the approximation factor are huge, making the algorithms impractical.

In this work, we seek algorithms with weaker but still useful approximation guarantees that are practical: simple, fast and space efficient. We introduce a class of algorithms called single pass algorithms. In such an algorithm we maintain a single pointer within each string, starting at the left. In each step, if the current symbols match we advance both pointers, otherwise we have a mismatch and choose one of the pointers to advance. Such an algorithm is specified by its advancement rule, which determines which pointer to advance. We consider particularly simple (possibly randomized) advancement rules where at each mismatch step the pointer advanced depends only on the number of mismatches seen so far and the randomness of the algorithm. It is easy to see that the total number of mismatches is always an upper bound on edit distance. Saha (2014) showed that the simple randomized rule (on mismatch advance a pointer at random) when run on two strings of edit distance d returns (with high probability) an upper bound of $O(d^2)$.

In this work we (1) present a deterministic single pass algorithm that achieves similar performance and (2) prove that no algorithm (even randomized) in this class can give a better approximation factor.

For the Dyck edit distance problem, Saha gave a complicated randomized reduction from Dyck edit distance to standard edit distance at a cost of a $O(\log d)$ factor where d is the Dyck edit distance. I will present a simple deterministic reduction with a similar (slightly better) approximation guarantee.

3.15 HDX Condensers

Amnon Ta-Shma (Tel Aviv University, IL)

Joint work of Itay Cohen, Ron Roth, Amnon Ta-Shma

More than twenty years ago, Capalbo, Reingold, Vadhan, and Wigderson gave the first (and up-to-date only) explicit construction of a bipartite expander with almost full combinatorial expansion. The construction incorporates zig-zag ideas and extractor technology and is rather complicated. We give an alternative construction that builds upon recent constructions of hyper-regular, high-dimensional expanders. The new construction is, in our opinion, simple and elegant.

Beyond demonstrating a new, surprising, and intriguing, application of high-dimensional expanders, the construction employs totally new ideas which we hope may lead to progress on the still remaining open problems in the area.

3.16 Cutting Planes Width and the Complexity of Graph Isomorphism Refutations

Jacobo Torán (Universität Ulm, DE)

Joint work of Jacobo Torán, Florian Wörz

The width complexity measure plays a central role in Resolution and other propositional proof systems like Polynomial Calculus (under the name of degree). The study of width lower bounds is the most extended method for proving size lower bounds, and it is known that for these systems, proofs with small width also imply the existence of proofs with small size. Not much has been studied, however, about the width parameter in the Cutting Planes (CP) proof system, a measure that was introduced by Dantchev and Martin in 2011 under the name of CP cutwidth.

In this talk, we consider the width complexity of CP refutations of graph isomorphism formulas. For a pair of non-isomorphic graphs G and H, we show a direct connection between the Weisfeiler–Leman differentiation number $\mathsf{WL}(G, H)$ of the graphs and the width of a CP refutation for the corresponding isomorphism formula $Iso(G, H)$. In particular, we show that if $\mathsf{WL}(G, H) \leq k$, then there is a CP refutation of $Iso(G, H)$ with width k, and if $\mathsf{WL}(G, H) > k$, then there are no CP refutations of $Iso(G, H)$ with width $k - 2$. Similar results are known for other proof systems, like Resolution, Sherali–Adams, or Polynomial Calculus. We also show polynomial-size CP refutations from our width bound for isomorphism formulas for graphs with constant WL.

3.17 Extractors for Algebraic Sources

Ben Lee Volk (Reichman University – Herzliya, IL)

Joint work of Zeyu Guo, Ben Lee Volk, Akhil Jalan, David Zuckerman
Main reference Zeyu Guo, Ben Lee Volk, Akhil Jalan, David Zuckerman: "Extractors for Images of Varieties", CoRR, Vol. abs/2211.14497, 2022.
URL https://doi.org//10.48550/arXiv.2211.14497

Randomness extractors are tools for converting "low quality" randomness into "high quality" randomness. In addition to being useful in the areas of pseudorandomness and derandomization, these objects are also connected to various fundamental notions in complexity theory and mathematics in general. In this talk we'll consider the randomness extraction problem from distributions with algebraic structure. We'll survey the different types of algebraic sources and constructions, and talk about a recent construction of extractors for polynomial images of varieties

4 Working groups

4.1 Lifting dichotomy theorems

Amit Chakrabarti (Dartmouth College – Hanover, US), Susanna de Rezende (Lund University, SE), Yuval Filmus (Technion – Haifa, IL), Mika Göös (EPFL Lausanne, CH), Johan Hastad (KTH Royal Institute of Technology – Stockholm, SE), Robert Robere (McGill University – Montréal, CA), and Avishay Tal (University of California – Berkeley, US)

Whenever there is a lifting theorem that works with constant size gadgets, there is hope to understand *all* gadgets. As a simple example, consider lifting decision tree depth to decision tree size. Using a simulation-type argument or a random restriction, it is not hard to check that $\log_2 \mathrm{DTsize}(f \circ \oplus_2) \geq \mathrm{DTdepth}(f)$. In fact, this works for any gadget g as long as g does not have certificates of size 1. If g does have a certificate of size 1 then up to negating inputs and outputs, g is either a (possibly degenerate) OR, or it projects to $g_0 = x \vee (y \wedge z)$. In the former case, depth does not lift to size (take f to be a large OR). In the latter case, we can lower bound $\log_2 \mathrm{DTsize}(f \circ g)$ by both the certificate complexity of f and (using a result of Sherstov) the degree of f; in particular, $\log_2 \mathrm{DTsize}(f \circ g) = \Omega(\mathrm{DTdepth}(f)^{1/2})$. We do not know whether the square root loss is necessary.

5 Open problems

5.1 Sampling modular distributions locally

Yuval Filmus (Technion – Haifa, IL)

Joint work of Yuval Filmus, Itai Leigh, Artur Riazanov, Dmitry Sokolov

Emanuele Viola initiated the study of the complexity of distributions. Given an infinite supply of iid unbiased random bits, which distributions can we sample in low complexity? Let us focus on locally samplable distributions. These are distributions such that for each $\epsilon > 0$ there is $d = d(\epsilon)$ and a d-local sampler (meaning that every output bit depends on at most d input bits) whose output is within variation distance ϵ of the target distribution.

The uniform distribution of all even parity strings is famously samplable with no error and locality 2. What about the uniform distribution over all strings whose Hamming weight is a multiple of m? We conjecture that for $m > 2$, this distribution cannot be sampled locally.

Participants

- Shyan Akmal
MIT – Cambridge, US
- Max Bannach
Universität zu Lübeck, DE
- Olaf Beyersdorff
Friedrich-Schiller-Universität Jena, DE
- Harry Buhrman
CWI – Amsterdam, NL
- Igor Carboni Oliveira
University of Warwick – Coventry, GB
- Gaia Carenini
ENS – Paris, FR
- Amit Chakrabarti
Dartmouth College – Hanover, US
- Sourav Chakraborty
Indian Statistical Institute – Kolkata, IN
- Gil Cohen
Tel Aviv University, IL
- Susanna de Rezende
Lund University, SE
- Yuval Filmus
Technion – Haifa, IL
- Anna Gál
University of Texas – Austin, US
- Mika Göös
EPFL Lausanne, CH
- Rohit Gurjar
Indian Institute of Technology – Mumbai, IN
- Kristoffer Arnsfelt Hansen
Aarhus University, DK
- Johan Hastad
KTH Royal Institute of Technology – Stockholm, SE
- Shuichi Hirahara
National Institute of Informatics – Tokyo, JP
- Rahul Ilango
MIT – Cambridge, US
- Swastik Kopparty
University of Toronto, CA
- Michal Koucký
Charles University – Prague, CZ
- Marvin Künnemann
RPTU – Kaiserslautern, DE
- Alexander S. Kulikov
JetBrains Research – Paphos, CY
- Sophie Laplante
Université Paris Cité, FR
- Zhenjian Lu
University of Oxford, GB
- Meena Mahajan
The Institute of Mathematical Sciences – Chennai, IN
- Jakob Nordström
University of Copenhagen, DK & Lund University, SE
- Manaswi Parashar
Aarhus University, DK
- Rüdiger Reischuk
Universität zu Lübeck, DE
- Robert Robere
McGill University – Montréal, CA
- Michael E. Saks
Rutgers University – Piscataway, US
- Rahul Santhanam
University of Oxford, GB
- Melanie Schmidt
Heinrich-Heine-Universität Düsseldorf, DE
- Amnon Ta-Shma
Tel Aviv University, IL
- Avishay Tal
University of California – Berkeley, US
- Till Tantau
Universität zu Lübeck, DE
- Thomas Thierauf
Hochschule Aalen, DE
- Jacobo Torán
Universität Ulm, DE
- Quinten Tupker
CWI – Amsterdam, NL
- Ben Lee Volk
Reichman University – Herzliya, IL