Review of Doctoral Thesis: "Limits of Data Structures, Communication, and Cards" by Mgr. Pavel Dvořák

Kristoffer Arnsfelt Hansen

Aarhus University
arnsfelt@cs.au.dk

1 Introduction

Pavel Dvořák's doctoral thesis "Limits of Data Structures, Communication, and Cards" is a contribution to computational complexity theory. The thesis is written in English, and it contains an introductory overview (4 pages), preliminaries on information theory (2 pages), and contents based on the following papers (chapters 2–5, 53 pages), selected amongst the extensive number of papers by author:

1. A. Chattopadhyay, P. Dvořák, M. Koucký, B. Loff, and S. Mukhopadhyay, "Lower Bounds for Elimination via Weak Regularity", STACS 2017, Leibniz International Proceedings in Informatics (LIPIcs) 66, 21:1-21:14, (2017).
2. P. Dvořák, and B. Loff, "Lower Bounds for Semi-adaptive Data Structures via Corruption", FSTTCS 2020, Leibniz International Proceedings in Informatics (LIPIcs) 182, 20:1–20:15, (2020).
3. P. Dvořák, M. Koucký, "Barrington Plays Cards: The Complexity of Card-Based Protocols", STACS 2021, Leibniz International Proceedings in Informatics (LIPIcs) 187, 26:1–26:17, (2021).
4. P. Dvořák, M. Koucký, K. Král, and V. Slívová, "Data Structures Lower Bounds and Popular Conjectures", appears only as a technical report, CoRR abs/2102.09294 (2021).

The numbering above is used in the evaluation below.

2 Evaluation

The introduction of the thesis provide a brief overview of several central topics of computational complexity, motivating the study of so-called conditional lower bounds, the study of restricted models of computation and their relationship, also enabling results for data structures. A large part of the mathematical definitions needed in thesis and review of prior work is deferred to subsequent chapters. A notable exception introduced separately is information theory which plays a central role in chapters 2–4.

Chapter 2, which is based on paper 2, is concerned with lower bounds for data structures. An approach to the outstanding open problem of proving polynomial lower bounds for dynamic data structures is the so-called multiphase problem defined proposed by Pătraşcu, which is a process consisting of consecutive phases of initialization, update, and query. A recent work of Ko and Weinstein studied data structures solving the multiphase problem but restricted to be *semi-adaptive* (which means that in the query phase, once the updates of the second phase are accessed, access to the original result of the initialization phase is prohibited) and obtained strong lower bounds in this setting.

The multiphase problem of Pătraşcu was defined in terms of the disjointness function and the results of Ko and Weinstein address only this case. However, a multiphase problem can be defined by any function. Dvořák shows how to extend the approach to give lower bounds for any function that has a large *smooth corruption* bound, involving a considerable novel technical contribution. This means that both the result and techniques of Ko and Weinstein are greatly clarified and several new (and stronger) lower bounds are obtained. It is plausible that the technical contribution involving in applying the smooth corruption bound may find additional uses in lower bounds.

Chapter 3, which is based on parts of paper 1, addresses the so-called *elimination problem* in communication complexity. Dvořák introduces a new complexity measure called *weak-regularity* (relaxing a notion of regularity used by Raz and Wigderson), and uses this to prove optimal lower bounds for the elimination problem of the greater than function GT. A more general result is proved in paper 1 based on discrepancy bounds, but the proof presented in this chapter is interesting in its own right and is developed from scratch in contrast to the discrepancy based proof that makes uses of an XOR lemma for discrepancy by Lee, Shraibman, and Špalek.

Chapter 4, which is based on paper 4, presents conditional lower bounds for data structures for solving the problems function inversion and polynomial evaluation/interpolation. The conditional aspect here is that the lower bounds are based on the so-called network coding conjecture (NCC), which recently were used to prove conditional lower bounds for integer multiplication and external memory integer sorting. Dvořák shows how to derive lower bounds for data structures supporting permutation inversion as well as polynomial evaluation and interpolation. The technique used is adapted from precious NCC based lower bounds together with some novel ideas.

The problem of function inversion is central in cryptography, where the security of many cryptographic systems rely on the computational difficulty of inverting certain functions, and the new results are interesting from this point of view also, even as they apply only to non-adaptive data structures. The problem of function evaluation and interpolation corresponds exactly to computing the Discrete Fourier Transform (DFT), here in the setting of finite fields. Computing the DFT is a fundamental computational task, solved by the Fast Fourier Transform, with a large number of applications, and the lower bounds are thus of fundamental interest.

Chapter 5, which is based on paper 3, is concerned with *card-based* protocols as a model of computation. Card-based protocols were introduced by Boer and has been studied in the cryptography community as a way for two parties to jointly compute a specific functions securely where the input of each player is hidden, and a substantial amount of research has been concerned with the study of card-based protocols. Dvořák presents several new results about card-based protocols from a computational complexity perspective. Using the classification of the complexity class NC^1 in terms of constant width permutation branching programs, the class of read-only oblivious protocols using a constant number of auxiliary cards are shown to also coincide with NC^1. For $s(n) \geq \log n$, any function computable in polynomial time and in space $O(s(n))$ on a Turing machine can be simulated by a card-based protocol with $O(s(n))$ auxiliary cards, which for $s(n) = O(\log n)$ leads to a precise characterization of L/poly. These results together fills a void in the literature. Dvořák also presents results about read-write protocols and more efficient encodings of inputs into cards.

3 Summary

The study of lower bounds in a central tropic in computational complexity of great importance to the study of algorithms and data-structures but also to neighboring areas such as cryptography.

Dvořák's thesis provides very interesting results in this direction. He proves new lower bounds for data-structures extending and improving the state-of-the-art of the field. He shows new results in communication complexity for the elimination problem. Finally he studies the model of card-based protocols and gives new and precise characterizations of the computational power of large classes of protocols.

The work provides a unified body of research representing significant progress in computational complexity theory. The work is of high quality and raises interesting questions for future research.

4 Conclusion

The work presented by Pavel Dvořák is of high quality, proving ability for creative scientific work.

August 2021

Kristoffer Arnsfelt Hansen