UNIVERSITÄT DES SAARLANDES

Saarland University
Postfach 15 11 50 | 66041 Saarbrücken | Germany

Instytut Matematyczny
pl. Grunwaldzki 2
50-384 Wrocław
Poland

Prof. Dr.-Ing. Karl Bringmann
Saarland University, Chair for Algorithms & Complexity

Saarland Informatics Campus
Building E 1 3
66123 Saarbrücken
Germany

T: +49 681 9325 1005
bringmann@cs.uni-saarland.de
www.uni-saarland.de

Date December 14, 2024

Subject Evaluation of Bartłomiej Dudek's PhD thesis "Equivalences between Some Problems on Strings, Trees and Graphs"

To whom it may concern,

In his PhD thesis, Bartłomiej Dudek makes fundamental contributions to several well-studied problems in theoretical computer science, and significantly advanced the field of fine-grained complexity theory. His two main contributions are an equivalence proof of 3SUM and all 3LDTs, solving a 20-year-old open problem, and an equivalence proof of counting 4-cycles, quartet distance, and counting 4-patterns in permutations. These problems have been studied for decades, and thus advancing the state of the art on any one of these problems would have been a great thesis on its own. Besides these main contributions, the thesis also presents an improved algorithm for online context-free grammar recognition, and a new algorithm for constructing top trees matching the information-theoretic lower bound. In total, this makes for an exceptional thesis. Let me discuss the contributions in more detail in what follows.

Equivalence of 3SUM and 3LDTs. In the 3SUM problem, given a set X of n numbers, the task is to decide whether for any numbers $a, b, c \in X$ we have $a + b + c = 0$. This problem can be solved in time $O(n^2)$ and is famously conjectured to require time $n^{2-o(1)}$. Many variations of 3SUM are known to be equivalent, e.g. variants with three given sets instead of one, asking for three distinct numbers or not necessarily distinct numbers, etc. Jeff Erickson asked 20 years ago whether the variant with $a + b = 2c$ was equivalent as well. This problem remained open, until Bartłomiej and his coauthors proved that it is equivalent to 3SUM at STOC'20. They even proved a vast generalization, namely for any integer constants $\alpha, \beta, \gamma, \delta$ the problem variant with $\alpha \cdot a + \beta \cdot b + \gamma \cdot c = \delta$ is equivalent to 3SUM (unless it is trivial because there exist no numbers $a, b, c \in \mathbb{N}$ with this property). The result is also very interesting on a technical level, because it involves Behrend sets, which had previously been used to prove hardness of Subset Sum and Bin Packing. Bartłomiej used Behrend sets for the first time on the polynomial-time problem 3SUM. It was quite surprising to me when I first saw this use of Behrend sets and how it leads to an answer to this 20-year-old question. Bartłomiej's work has also inspired interesting follow-up work, for example he left open 4LDTs, in particular whether they all require quadratic time, which was later resolved by Jin and Xu who proved that all 4LDTs are 3SUM-hard.

Equivalence of counting 4-cycles, quartet distance, and counting 4-patterns in permutations. In this part Bartłomiej proves that three seemingly completely unrelated problems are equivalent, more precisely that they have the same time complexity up to logarithmic factors. The first problem is counting the number of 4-cycles in a given m-edge graph, which was known to be in time $O(m^{1.48})$. The second problem is the so-called quartet distance, which is a distance measure on two trees with the same leaves, which counts the number of 4-tuples of leaves that have a different topology in the two trees. This is seemingly a completely different problem that counting 4-cycles in graphs. However, via deep combinatorial insights into the problem structure and a surprising construction, Bartłomiej and his advisor found a way to connect the two problems in a tight way, proving that they have the same time complexity up to logarithmic factors. This allowed them to improve the known $\tilde{O}(n^2)$-time algorithm for quartet distance to time $O(n^{1.48})$. Moreover, the equivalence means that any future progress on one of the two problems automatically carries over to the other, and it shows that both problems are hard for the same reason. In this sense, the quartet distance can now be considered resolved, because it can be fully explained by pointing to the complexity of another, even more fundamental problem. In addition to the running time in terms of the input size n, Bartłomiej also achieved improvements in terms of the value of the quartet distance d, specifically he improved the known running time of $\tilde{O}(nd)$ to $\tilde{O}(nd^{0.69})$ and $\tilde{O}(n^{1.16}d^{0.43})$. These results were presented at STOC'19.
The third problem is counting 4-patterns in permutations. Again this is a seemingly completely unrelated problem, but Bartłomiej managed to prove an equivalence with counting 4-cycles, again fully explaining the time complexity of one problem in terms of another. This also results in improved algorithms for counting 4-patterns in permutations. These results were presented at ISAAC'20, were they received a best paper award.

Online Context-Free Grammar (CFG) Recognition. In CFG Recognition, given a context-free grammar of constant size and a string of length n, the task is to decide whether the grammar accepts the string. Valiant famously solved this problem in time $O(n^{\omega})$, where ω is the exponent of fast matrix multiplication. Bartłomiej studied CFG Recognition in a well-motivated online setting, where the string is revealed character by character, and at any point in time the task is to decide whether the current prefix is accepted by the CFG. Since CFG Recognition is connected to matrix multiplication by Valiant's algorithm, it is natural to ask whether Online CFG Recognition is connected to the OMv problem, an online variant of matrix multiplication that is widely used for conditional lower bounds in fine-grained complexity theory. Indeed, Bartłomiej found such a connection, by an interesting adaptation of Valiant's algorithm. Combined with the state-of-the-art algorithms for OMv, he obtains an improved algorithm for Online CFG Recognition. This result was presented at CPM'24.

Constructing Top Trees. One of the technical ingredients that Bartłomiej used for his equivalence of counting 4-cycles and quartet distance are so-called top trees. As a side result, he presents new insights on top trees. He shows that a well-known construction of top trees is suboptimal, as it is a loglog-factor away from the information-theoretic lower bound, which answers an open problem from the literature on the optimality of this construction. Then he presents a slight variation of this well-known construction and proves that his variation matches the information-theoretic lower bounds. These results were presented at CPM'18 (and published in the journal TCS).

In summary, Bartłomiej Dudek's PhD thesis makes fundamental contributions to theoretical computer science, by finding tight connections between seemingly unrelated problems and by solving a 20-year-old open problem in the field of fine-grained complexity theory. He presents deep combinatorial insights and demonstrates great knowledge and technical skills in his surprising applications of non-trivial tools such as Behrend sets and top trees. The publications venues are outstanding: he published two papers at STOC and received a best paper award from ISAAC. This is an exceptional thesis, that clearly merits a PhD degree.

Karl Bringmann

Prof. Dr.-Ing. Karl Bringmann
Fachrichtung Informatik
Universität des Saarlandes
Saarland Informatics Campus
Gebäude E1 3
D-66123 Saarbrücken