Abstract

Traditionally, in the complexity theory, *easy* and *hard* problems are defined with respect to their polynomial time solvability. However, from the point of view of applicability in real-life applications, say biological or other big data, such an approach is too coarse. In particular, for some *easy* problems, the asymptotically fastest algorithms have polynomial time complexity with degree of the polynomial being quite high. This raises a natural question about the underlying difficulty of these problems:

Can we hope to improve these algorithms, or is there a fundamental reason for the lack of progress?

In this thesis we consider various problems on strings, trees and graphs. For some of them we show efficient algorithms, and in many cases, we also provide conditional lower bounds. That research falls within the area of fine-grained complexity whose goal is to create a map of polynomial-time problems by understanding the relationships between them.

Our first main contribution is establishing the subquadratic equivalence between all variants of the 3LDT problem, which generalizes the well-known 3SUM problem. In 3SUM, the goal is to find three elements from a given set that sum to zero. We show, that for all, so-called, non-trivial coefficients $\alpha_1, \alpha_2, \alpha_3$ and t, detecting whether a set contains three elements x_1, x_2, x_3 such that $\sum_i \alpha_i x_i = t$ is equivalent 3SUM. This also answers an open question posed by Jeff Erickson over 20 years ago regarding the difficulty of detecting three numbers in a set that form an arithmetic progression.

Our second contribution reveals that the following three seemingly unrelated problems are in fact equivalent: counting 4-cycles in a graph, computing the quartet distance between two trees and counting 4-patterns in permutations. This equivalence implies that all these problems can be solved with the same time complexity (up to polylogarithmic factors), currently $\mathcal{O}(n^{1.48})$. For the latter two problems, this allows us to improve the state-of-the-art algorithms, while also identifying a common barrier to reducing their complexity below $\mathcal{O}(n^{4/3})$.

The most technically challenging reduction in this thesis connects computing the quartet distance to counting 4-cycles, which, among other techniques, utilizes the top tree decomposition. To introduce this concept we also show that slowing down a top tree compression algorithm might lead to a better compression ratio.

As a final contribution, we reduce the Online Context-Free Recognition problem to multiple instances of Online Matrix-Vector Multiplication. This gives us an $n^3/2^{\Omega(\sqrt{\log n})}$ algorithm, which is the first improvement for this problem since 1985.