Introduction: ACM-SIAM Symposium on Discrete Algorithms (SODA) 2022 Special Issue

We are delighted to present a Special Issue of *ACM Transactions on Algorithms*, containing full versions of six articles that were presented at the 31st Annual ACM-SIAM **Symposium on Discrete Algorithms (SODA)** 2022 in Salt Lake City, UT, USA, on 5–8 January 2022. These articles, selected on the basis of their high rating by the conference program committee, have been thoroughly reviewed according to the journal's highest standards. We summarize their contributions below.

Multistage stochastic integer programs are a fundamental class of block structured integer programs which model optimization problems over a scenario tree. While these models are in general NP-hard, there has been tremendous progress over the last decade in the development of fixed parameter algorithms when the number of decision stages, the maximum size of a decision block, and the maximum size of any coefficient in the constraint matrix are fixed. For related block structures, in particular if the constraint matrix is the transpose of a multistage matrix, algorithms with nearly optimal parameter dependencies were obtained under the exponential time hypothesis. However, for multistage IPs, the best upper bound had a tower-like dependence on the parameters, while the best lower bound was doubly exponential (even for two-stage IPs). The article by Kim-Manuel Klein and Janina Reuter, "Collapsing the Tower—On the Complexity of Multistage Stochastic IPs" (https://doi.org/10.1145/3604554), closes this gap by improving the upper bound to doubly exponential when the number of stages is fixed.

The greedy spanner is a basic primitive that builds a sparse subgraph of a weighted undirected graph that approximately preserves pairwise distances. Given a desired stretch $t \geq 1$, the greedy spanner considers edges of the graph in non-decreasing order of weight and adds the edge if the shortest path distance in the current spanner is greater than t times the shortest path distance in the full graph. While these spanners are typically quite sparse, this is not necessarily enough to enable fast algorithms. For the sake of efficiency, a more useful property is the existence of sublinear sized balanced vertex separators, which have been used extensively to design fast recursive algorithms for problems on graphs embedded in a Euclidean space (the distance between any two vertices is the Euclidean distance). In this context, it was known that in two dimensions all subgraphs of a greedy $(1 + \epsilon)$-spanner have sublinear separators; however, it was open whereas this property extends to higher dimensions. The article by Hung Le and Cuong Than, "Greedy Spanners in Euclidean Spaces Admit Sublinear Separators" (https://doi.org/10.1145/3590771), fully resolves this question. Furthermore, the authors give a simple sufficient geometric criterion for establishing the existence of sublinear sized separators, relating to the maximum number of edges of length r that can be cut by a ball of radius r, which in particular holds for the greedy spanner.

ACM Reference format:

Daniel Dadush, Martin Milanič, and Tami Tamir. 2024. Introduction: ACM-SIAM Symposium on Discrete Algorithms (SODA) 2022 Special Issue. *ACM Trans. Algor.* 20, 3, Article 21 (June 2024), 2 pages.
https://doi.org/10.1145/3655622

ACM 1549-6333/2024/6-ART21
https://doi.org/10.1145/3655622

Hopcroft's problem, first posed by John Hopcroft in the early 1980s, is the following basic computational geometry problem: given n points and n lines in 2D, detect (or count, or report all) point-line incidence pairs. The article by Timothy M. Chan and Da Wei Zheng, titled "Hopcroft's Problem, Log-Star Shaving, 2D Fractional Cascading, and Decision Trees" (https://doi.org/10.1145/3591357), shows that Hopcroft's problem can be solved in time $O(n^{4/3})$, matching the conjectured lower bound and improving the best previous time bound obtained almost 30 years ago by Matoušek. Two interesting and different ways to achieve the result are described, one of which extends to any constant dimension. These new ideas have many consequences for related fundamental problems about geometric range searching.

Determining the computational complexity of the graph isomorphism problem is a long-standing open question in theoretical computer science. In the article titled "Isomorphism Testing for Graphs Excluding Small Topological Subgraphs" (https://doi.org/10.1145/3651986), Daniel Neuen addresses this question by giving an isomorphism test that runs in time $n^{\text{polylog}(h)}$ on all n-vertex graphs excluding some h-vertex graph as a topological subgraph. This result improves on the previous bounds of $n^{\text{polylog}(n)}$ (Babai, STOC 2016) and $n^{f(h)}$ for some function f (Grohe and Marx, *SIAM J. Comp.*, 2015) and unifies and extends previous isomorphism tests for graphs of maximum degree d running in time $n^{\text{polylog}(d)}$ (Grohe, Neuen, and Schweitzer, FOCS 2018) and for graphs of Hadwiger number h running in time $n^{\text{polylog}(h)}$ (Grohe, Wiebking, and Neuen, FOCS 2020).

A Dyck sequence is a sequence of opening and closing parentheses (of various types) that is balanced. The Dyck edit distance of a given sequence of parentheses S is the smallest number of edit operations needed to transform S into a Dyck sequence. In the article titled "An Improved Algorithm for the k-Dyck Edit Distance Problem" (https://doi.org/10.1145/3627539), Dvir Fried, Shay Golan, Tomasz Kociumaka, Tsvi Kopelowitz, Ely Porat, and Tatiana Starikovskaya present new algorithms for computing the Dyck distance of sequences given a threshold k. Specifically, the input is a sequence of parentheses, S, and a positive integer k, and the goal is to compute the Dyck edit distance of S only if the distance is at most k, and otherwise report that the distance is larger than k. The authors improve the best known bound for this problem for both deterministic and randomized algorithms.

In the article titled "Better Sum Estimation via Weighted Sampling" (https://doi.org/10.1145/3650030), Lorenzo Beretta and Jakub Tětek studied a setting in which there are a number of weighted items, and we are interested in estimating their total weight to be within factor of $1 \pm \epsilon$, with some constant probability $> \frac{1}{2}$. This problem has applications in the area of sublinear-time graph algorithms. In their setting, items may be sampled uniformly, or with probabilities proportional to their weights, or in both ways (hybrid setting). The article improves both upper and lower bounds known for the proportional and the hybrid setting and also provides the first results for the previously unexplored scenario in which n is not known in advance.

Daniel Dadush
CWI and Utrecht University

Martin Milanič
University of Primorska

Tami Tamir
Reichman University

SODA 2022 Special Issue Guest Editors