An Optimal Algorithm for Sorting Pattern-Avoiding Sequences

Michal Opler Czech Technical University in Prague

A standard argument shows that any comparison-based algorithm for sorting a set of permutations Γ, each of length n, must make at least $\Omega(\log|\Gamma| + n)$ comparisons in the worst case. In order to see how this lower bound applies to pattern-avoiding sequences, it invites the natural question of how many pattern-avoiding permutations of given length are there. In fact, it has been a long-standing open problem whether the number of π-avoiding permutations of length n grows at most exponentially in n. This question, known as the *Stanley-Wilf conjecture*, was eventually answered positively by Marcus and Tardos [8] building on the works of Klazar [7] and Füredi and Hajnal [6]. In fact, the limit $s_\pi = \lim_{n\to\infty} \sqrt[n]{|\operatorname{Av}_n(\pi)|}$ exists [1], where $\operatorname{Av}_n(\pi)$ denotes the set of all π-avoiding permutations of length n; s_π is known as the *Stanley-Wilf limit* of π. It follows that any comparison-based algorithm for sorting π-avoiding inputs must make at least $\Omega((\log s_\pi + 1)\cdot n)$ comparisons in the worst case.

Our main result is an asymptotically optimal sorting algorithm for pattern-avoiding sequences that matches this information-theoretic lower bound for each π up to a global constant. Furthermore, the algorithm requires prior knowledge of neither the forbidden pattern π nor its Stanley-Wilf limit s_π.

Theorem 1. *There is an algorithm that sorts a π-avoiding sequence of length n in $O((\log s_\pi + 1)\cdot n)$ time even if π is a priori unknown.*

The algorithm is built from two components. The first component is an efficient multi-way merge procedure capable of merging up to $\frac{n}{\log n}$ sequences with n elements in total in linear time, as long as the sequences originate from a π-avoiding input. Its analysis is heavily based on the Marcus-Tardos theorem [8].

The second component is an efficient algorithm to sort a large set of short sequences. Naturally, many sequences in such a set must be pairwise order-isomorphic, which suggests that we could try grouping them into equivalence classes and then actually sorting only one sequence from each class. Unfortunately, computing the equivalence classes is in some sense as hard as sorting itself. Instead, we take a different approach of guessing a shallow decision tree for sorting π-avoiding permutations guaranteed by the result of Fredman [5].

Theorem 1 then follows by first cutting up the input into sequences of length roughly $\log\log\log n$ which we can efficiently sort using the optimal decision tree. The sorted sequences then serve as a basis for a bottom-up merge sort using the efficient multi-way merge.

Furthermore, let us point out that Theorem 1 achieves the information-theoretic lower bound for arbitrary π-avoiding sequences, not only permutations. This is in sharp contrast with previous works [3, 2] where the generalization beyond permutations comes with a price of worsening the dependence on π. In order to achieve this, we strengthen a combinatorial result of Cibulka [4] and show that the number of π-avoiding $n\times n$ binary matrices with n 1-entries is at most $c_\pi^{2n+O(1)}$.

References

[1] R. Arratia, "On the Stanley-Wilf conjecture for the number of permutations avoiding a given pattern," *Electron. J. Comb.*, vol. 6, 1999.

[2] B. A. Berendsohn, L. Kozma, and M. Opler, "Optimization with pattern-avoiding input," in *56th Annual ACM Symposium on Theory of Computing, STOC 2024*. ACM, 2024, pp. 671–682.

[3] P. Chalermsook, M. Goswami, L. Kozma, K. Mehlhorn, and T. Saranurak, "Pattern-avoiding access in binary search trees," in *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015*. IEEE Computer Society, 2015, pp. 410–423.

[4] J. Cibulka, "On constants in the Füredi–Hajnal and the Stanley–Wilf conjecture," *Journal of Combinatorial Theory, Series A*, vol. 116, no. 2, pp. 290–302, 2009.

[5] M. L. Fredman, "How good is the information theory bound in sorting?" *Theor. Comput. Sci.*, vol. 1, no. 4, pp. 355–361, 1976.

[6] Z. Füredi and P. Hajnal, "Davenport-Schinzel theory of matrices," *Discret. Math.*, vol. 103, no. 3, pp. 233–251, 1992.

[7] M. Klazar, "The Füredi-Hajnal Conjecture Implies the Stanley-Wilf Conjecture," in *Formal Power Series and Algebraic Combinatorics: 12th International Conference, FPSAC'00*. Springer, 2000, pp. 250–255.

[8] A. Marcus and G. Tardos, "Excluded permutation matrices and the Stanley–Wilf conjecture," *Journal of Combinatorial Theory, Series A*, vol. 107, no. 1, pp. 153–160, 2004.