GRWC 2019 Problem Statement

For the 2019 Rocky Mountain-Great Plains Graduate Research Workshop in Combinatorics at The University of Kansas, I propose Question 1.

1. Definitions

A *permutation of* $[n]$ is an string $\sigma_1 \cdots \sigma_n$ consisting of each element from $[n] = \{1, \dots, n\}$ exactly once. We write $\mathcal{S}_n$ for the space of all permutations of length n. Given two permutations $\sigma, \tau \in \mathcal{S}_n$, we define their *Kendall tau* or *bubble-sort distance* to be $D_1(\sigma, \tau)$, the number of relative inversions, i.e., pairs $\{i, j\}$ of indices such that σ and τ disagree on the relative sizes of the corresponding entries. Equivalently, $D_1(\sigma, \tau)$ is the number of "adjacent swaps" needed to transform σ into τ. For convenience, we normalize $D_1(\sigma, \tau)$ as $d_1(G, H) := D_1(\sigma, \tau)/\binom{n}{2}$ so that $0 \leq d_1 \leq 1$.

Given permutations $\pi \in \mathcal{S}_k$ and $\sigma \in \mathcal{S}_n$ (typically with $k \leq n$), we say that σ *forbids* μ if there does not exist a subword $\sigma_{i_1} \cdots \sigma_{i_k}$ of σ so that $\pi_\ell < \pi_m$ if and only if $\sigma_{i_\ell} < \sigma_{i_m}$. Further, we write $\mathrm{Forb}_n(P)$ for the set of all permutations on $[n]$ which forbid each $\pi \in P$. We also suppress the n subscript when the length is not restricted.

A *graph on* $[n]$ is the pair $([n], E)$ with $[n]$ the *vertex set* and $E \subseteq \binom{[n]}{2}$ a subset of unordered pairs $ij = \{i, j\}$, known as *edges*. We write $\mathcal{G}_n$ for the set of all graphs on $[n]$. Given two graphs $G, H \in \mathcal{G}_n$, we define their *edit distance* to be $D_1(G, H) := |E(G) \triangle E(H)|$, the cardinality of the symmetric difference of their edge sets. Equivalently, $D_1(G, H)$ is the number of edge removals and additions needed to transform G into H. For convenience, we normalize $D_1(G, H)$ as $d_1(G, H) := D_1(G, H)/\binom{n}{2}$ so that $0 \leq d_1 \leq 1$.

Given graphs $F \in \mathcal{G}_k$ and $G \in \mathcal{G}_n$, we say that G *forbids* F if there does not exist an induced subgraph of G isomorphic to F. Further, we write $\mathrm{Forb}_n(\mathcal{F})$ for the set of all graphs on $[n]$ which forbid each $F \in \mathcal{F}$. For convenience, we suppress the n subscript when the number of vertices is not restricted.
Note: Graph properties of the form $\mathrm{Forb}(\mathcal{F})$ are commonly known as *hereditary properties.*

2. Overview

Question 1. *What permutations are the furthest from avoiding some permutations P? In other words, (a) what is*

$$d_{1,n}(P) := \max_{\sigma \in \mathcal{S}_n} d_1(\sigma, Forb_n(P)) = \max_{\sigma \in \mathcal{S}_n} \min_{\tau \in Forb_n(P)} d_1(\sigma, \tau),$$

(b) does $\lim_{n\to\infty} d_{1,n}(P)$ exist, and (c) what permutations σ achieve the maximum?

Question 1 is motivated by the so-called *edit distance problem for hereditary properties* which asks the analogous problem for graphs. In [1] Alon and Stav addressed (a–c) for graphs by showing that for any graph property of the form $\mathrm{Forb}(\mathcal{F})$, there exists some $0 \leq p_0 \leq 1$ so that the Erdős-Rényi random graph model $\mathbb{G}(n, p_0)$ generates asymptotically maximizers, almost surely.

Theorem 2. *For all families $\mathcal{F}$ of graphs, there exists some $0 \leq p_0 \leq 1$ so that*

$$\lim_{n\to\infty} \max_{G \in \mathcal{G}_n} d_1(G, Forb_n(\mathcal{F})) = \lim_{n\to\infty} \mathbb{E}_{G \sim \mathbb{G}(n,p_0)} \left[d_1(G, Forb_n(\mathcal{F})\right] .$$

Here, $G \sim \mathbb{G}(n, p)$ means that G is generated by indepently deciding for each possible edge ij that $ij \in E(G)$ with probability p. In [2] Balogh and Martin improved Theorem 2 by showing that for each $0 \leq p \leq 1$, the random graph model $\mathbb{G}(n, p)$ asymptotically generates the furthest graphs of density p from $\text{Forb}(\mathcal{F})$, almost surely. See [3] for a more detailed discussion. In fact, graphons, a tool developed in the emergent field of combinatorial limit theory, may be used to prove both Theorem 2 and the Balogn and Martin generalization using simple geometric arguments.

Since $\mathbb{G}(n,p)$ generates random graphs of edge density concentrated around p, should it also be the case that the permutations σ which maximize $d_1(\sigma, \text{Forb}_n(P))$ are "random-like" in some sense? Further, are the maximizers among the set of permutations with inversion density p also "random-like"? A satisfying answer to Question 1 may suggest a heuristically good analogue of $\mathbb{G}(n,p)$ (for $p \neq 1/2$) in the setting of permutations, which permutation theory arguably lacks.

As was the case with the edit distance problem, the problem of determing an answer to Question 1.(a) is of computational interest. We ask the following.

Question 3. *What permutations are furthest from Forb$(1 \cdots k)$? What permutations are furthest from Forb(132)? How far are they?*

Elements of $\text{Forb}(1 \cdots k)$ are precisely an increasing concatenation of at most $k-1$ decreasing intervals and 132-free permutations are precisely "triangle-shaped", characterized by the overlay of a monotone decreasing subword and a monone increasing subword.

3. Plan

For the first week, I suspect that two groups will form, each to address one of Questions 1 and 3. Due to the classification of $\text{Forb}(1 \cdots k)$ and $\text{Forb}(132)$, I suspect computation of at least $\text{Forb}(123)$ and $\text{Forb}(132)$ are attainable in that week.

If interest continues through the second week and Question 3 is mostly solved, we could move on the Balogh Martin-esque generalization of Question 3 which seeks optimizers among the set of inversion density p permutations.

Question 1 does not appear hopeless to me, but would likely require more time. Permutations have their own regularity lemma and their own combinatorial limit object known as a permuton. As was the case with graph limits, it may be possible to make another simple geometric argument.

References

[1] Noga Alon and Uri Stav, *What is the furthest graph from a hereditary property?*, Random Structures Algo- rithms **33** (2008), no. 1, 87–104. MR 2428979 (2009c:05219)

[2] Jószef Balogh and Ryan R. Martin *Edit distance and its computation*, Electron. J. Combin. **15** (2008), no. 1, Research Paper 20, 27. MR 2383440 (2008j:05175)

[3] Ryan R. Martin, The edit distance in graphs: methods, results, and generalizations. *Recent trends in combinatorics*, 31–62, IMA Vol. Math. Appl., 159, *Springer, [Cham]*, 2016.