Compact dot representations in permutation avoidance

Arjun Khandelwal

mentor Rik Sengupta

PRIMES final paper
draft

March 3, 2015

0.1 Introduction

We say a permutation π contains a permutation π' if there exists a subsequence of π order-isomorphic to π'. The study of permutation avoidance (and, resultingly, permutation containment) can be traced back to Donald Knuth's work on stack-sortable permutations in the late 1960s. The "dual" to the notion permutation avoidance is that of pattern-packing, or the study of permutations which contain the largest number of smaller permutations. Of particular interest is the problem of the smallest "superpattern", or the smallest permutation which contains all permutations of a given length, originally suggested by Arratia. The trivial lower bound for super-pattern length is $\frac{k^2}{e^2}$ for a k-permutation, and a construction of Miller gave an upper bound of $\frac{k(k+1)}{2}$. Researchers disagree on the asymptotic limit of the minimal length of a superpattern - some estimate the lower bound is tight, whereas others conjecture a lower bound of $\frac{k^2}{4}$ or $\frac{k^2}{2}$. A paper published by a Eriksson et. al in 2001 introduced a new form of representing a permutation, referred to as the compact dot representation, with the goal of constructing a smaller superpattern. We study this representation and give bounds on its size. We also consider a variant of the problem, where limitations on the alphabet size are imposed, and obtain lower bounds. Lastly, we consider the Mobius function of the poset of permutations ordered by containment.

0.2 Compact Dot Representations

In [1], a new method of representing a permutation was introduced relating to its position in a "tilted checkerboard". This representation was used to improve the standing upper bound on the minimal length super-pattern, in addition to being an object of interest in its own right.

0.2.1 Definitions

Definition An *n-permutation* is an ordering of the indices $1, 2, 3, ..., n$.

Definition An *ascent* in a permutation π occurs between two consecutive indices i and $i + 1$ if $\pi(i) < \pi(i + 1)$.

Definition A *descent* in π occurs between two consecutive indices i and $i+1$ if $\pi(i) > \pi(i + 1)$.

Definition An *inverse descent* in π occurs wherever there is a descent in the inverse permutation of π.

Definition The *n-tilted square* is the n^2 permutation obtained by ordering the indices $1, 2, ..., n^2$ into n sequences of length n each decreasing by n and putting the sequences in ascending order [Eriksson et. al]

Example The "4-tilted square" corresponds to the permutation 13 9 5 1 14 10 6 2 15 11 7 3 16 12 8 4.

Definition The *compact dot representation* of a permutation is the set of points representing the permutation in the grid of the tilted square farthest to the west and south.

Definition Let $A(\pi)$, $D(\pi)$, $a(\pi)$, $d(\pi)$ denote the number of ascents, number of descents, number of inverse ascents, and number of inverse descents, respectively in a permutation π.

Lemma 0.2.1. *The compact dot representation of a permutation is achieved by taking each symbol π_i in π and placing a dot in the site with as many columns to the left of it as there are ascents before it and as many rows below it as there are inverse descents under it.*

The statement above was shown in [Eri13]. Note that we may thus bound representations under lines of the form $A(\pi)+d(\pi) = c$, where $A(\pi)$ and $d(\pi)$ denote the number of ascents and inverse descents of π, respectively. This will be the goal of the following theorem.

Theorem 0.2.2. *Let T_i denote the i-th triangular number. If $T_{i-1} < n \leq T_i$, then $A(\pi) + d(\pi) \leq 2n - 2i + 2$. Furthermore, this bound is tight.*

Proof. Note that $A(\pi) + D(\pi) = n$ and $a(\pi) + d(\pi) = n$ trivially for any n-permutation σ. Adding, we have

$$A(\pi) + d(\pi) + a(\pi) + D(\pi) = 2n$$

Furthermore,

$$a(\pi) + D(\pi) = A(\pi^C) + d(\pi^C)$$

, where π^C denotes the n-permutation obtained by replacing each symbol π_i with $n+1-\pi_i$, as (inverse) ascents become (inverse) descents (and vice-versa) when taking the complement. Thus, to maximize $A(\pi) + d(\pi)$ we determine the minimum of $A(\pi^C) + d(\pi^C)$. As π^C ranges over all n-permutations, we must to assign each of the n points to a unique square in the grid of the tilted square. This is equivalent to selecting n distinct lattice points with non-negative integer coordinates while trying to minimize the largest taxicab distance between the origin and any of the n points. This problem is solved by progressively selecting points along the line $x + y = c$, with steadily growing c. As there are n points to fill, we must have $c \geq i$ where $T_{i-1} < n \leq T_i$. Note that $c = i$ precisely when the arrangement of points includes n points inside and on the boundary of the triangle with vertices $(0, 0)$, $(0, i-1)$, and $(i-1, 0)$, which, reverting to the original problem, gives us $A(\pi) + d(\pi) \geq 2i - 2 \implies A(\pi) + d(\pi) \leq 2n - 2i + 2$. This bound is tight, as demonstrated by the permutation (for $n = 6$) $\pi^C = 6\ 4\ 1\ 5\ 2\ 3$ and $\pi = 1\ 3\ 6\ 2\ 5\ 4$. □

0.3 Superpatterns with restrictions on alphabet size

A natural extension to the superpattern problem, first considered by [2], asks the minimum length of a superpattern on a restricted alphabet. In this instance, repetition in the superpattern is allowed to occur.

Definition $n(l, m)$ denotes the length of the smallest pattern which contains every permutation of length m on at most l letters.

Example $n(4, 3) = 6$, as shown by the pattern 134213

Theorem 0.3.1. $n(m + c, m) \geq \frac{n(m,m)}{c+1}$ *for* $c < m$

Proof. Let π be an arbitrary superpattern of length l on the alphabet of size $m + c$, represented by $\pi_1\pi_2...\pi_l$. For each number $i \in [1, m + c]$, we let B_i be the sequence with symbols in $[1, m]$ defined by $i - ci - c + 1...i$ (removing all symbols outside the interval $[1, n]$ in the process, so for example B_1 and B_{m+c-2} would be the sequences 1 and $m - 2m - 1m$. We claim the sequence $B_{\pi_1}B_{\pi_2}...B_{\pi_l}$ of length less than or equal to $l(c + 1$ is a superpattern (which we denote ρ). Suppose an m-permutation ω is contained in π but not in ρ. Let $x_1, x_2, ..., x_m$ be the symbols in π corresponding to an occurrence of ω. Then, for some i, B_{x_i} must not contain ω_i.

Case 1: ω_i is in $(x_i, m]$

It follows that there are $\omega_i - 1$ symbols of ω (namely, $1, 2, ..., \omega_i - 1$) that must correspond by some of the symbols $1, 2, ..., x_i - 1$ in π. However, as $\omega_i > x_i$, we have $\omega_i - 1 > x_i - 1$. By the pigeonhole principle, at least one of the symbols $1, 2, ..., x_i - 1$ must correspond to multiple symbols in $1, 2, ..., \omega_i - 1$, which is a contradiction.

Case 2: ω_i is in $[1, x_i - c)$

It follows that there are $m - \omega_i$ symbols of ω (namely, $\omega_i + 1, \omega_i + 2, ..., m$) that must correspond to some of the symbols $x_i + 1, x_i + 2, ..., m + c$. As

$\omega_i < x_i - c$, we have $m - \omega_i > m + c - x_i$. By the pigeonhole principle, at least one of the symbols $x_i + 1, x_i + 2, ..., m + c$ must correspond to multiple symbols in $\omega_i + 1, \omega_i + 2, ..., m$, which is a contradiction.

Thus, ρ is a superpattern, from which the theorem follows immediately. □

Using the lower bound of $dn^{7/4+\epsilon}$ for $n(m, m)$ given in [3], where $\epsilon > 0$ and d is a constant dependent on ϵ, we arrive at the following.

Corollary 0.3.2. *$n(m + c, m) \geq \frac{n^2 - dn^{7/4+\epsilon}}{c+1}$, where $\epsilon > 0$ and d is a constant dependent on ϵ.*

It is interesting to note that for $c = 1$, the bound is asymptotically tight, as in [4] it was shown that there exists a superpattern of length $\frac{n(n+1)}{2}$ on the alphabet of size $m+1$ containing all permutations of length m. An interesting (and more approachable when compared to the general case) direction of study would be to tighten bounds $n(l, m)$ for other small c. Miller conjectures that for all l, $n(l, m) \geq \frac{n^2}{2}$, giving a wide gap between the bounds beginning at $c = 2$.

0.4 The Möbius function of the permutation pattern poset

Let P be the graded poset of permutations ordered by containment. That is, for arbitrary permutations π and σ, $\pi \leq \sigma$ if and only if σ contains π. This poset has been extensively studied by researchers, with increasing recent focus on its the Möbius function and topology. Here, after fixing a permutation π, we give an infinite non-trivial family of permutations pi' such that $\forall \omega \in \pi', \mu(\pi, \omega) = 0$. After concluding the proof, we found that this result follows from a result in [5].

Definition The Mobius function of an interval $[\sigma,\pi]$, denoted by $\mu(\sigma,\pi)$, is defined piece-wise as follows.

$$\mu(\sigma,\pi) = \begin{cases} 1 & : \sigma = \pi \\ 0 & : \sigma \not\leq \pi \\ -\sum\limits_{z\in[\sigma,\pi)} \mu(\sigma,z) & : \sigma < \pi \end{cases}$$

Note that

$$\sum_{z\in[\sigma,\pi]} \mu(\sigma,z) = 0$$

This is of fundamental importance in computing the Möbius function of arbitrary intervals.

Definition The family π' of permutations a with respect to a permutation π and a specified k-permutation τ contained exactly once in π may by defined in the following manner. First fix an integer $c \geq 2$. Let the indices corresponding to the occurence of τ in π be $a_1, a_2, ..., a_k$. We select an arbitrary a_i, and add c to all letters $\pi(n)$ of π such that $\pi(n) > \pi(a_i)$. Then we replace the symbol $\pi(a_i)$ in the resulting sequence with the sequence $\pi(a_i), \pi(a_i)+1, ..., \pi(a_i)+c$ if $\pi(a_i) > \pi(a_{i+1})$ and $\pi(a_i)+c, \pi(a_i)+c-1, ..., \pi(a_i)$ if $\pi(a_i) < \pi(a_{i+1})$ (we refer to these sections as A in the following proof). The resulting permutations (dependent on our choice of c and a_i) together make up π'.

Theorem 0.4.1. *Let τ be an arbitrary k-permutation, and let π be a permutation which contains exactly one occurence of τ. Then, for all permutations $\theta \in \pi'$, $\mu(\tau,\theta) = 0$.*

Proof. We induct on the length n of π. The base case occurs when $n = k$ (when $\pi = \tau$). Let ρ be an arbitrary permutation in π'. Then, every permutation z with $\tau \leq z < \rho$ must delete at least one symbol in A, as any permutation obtained by deleting a symbol outside of A would then not contain τ. Consider the permutation ρ' obtained from ρ by solely deleting one letter from A - this permutation contains all z and is also strictly contained in ρ. Thus, we have

$$\mu(\tau, \rho) = - \sum_{z \in [\tau, \rho)} \mu(\tau, z) = \sum_{z \in [\tau, \rho']} \mu(\tau, z) = 0$$

Assume the result holds for all $n < l$. Then, $z \in [\tau, \rho)$ may be separated into two categories; those that are obtained by deleting symbols in A (with potentially some deletions outside A as well) and those that are obtained by solely deleting symbols outside of A. Permutations γ in the latter category have $\mu(\tau, \gamma) = 0$ by the inductive hypothesis; thus we need not consider these permutations in the summation when computing $\mu(\tau, \rho)$. Here again, consider the permutation ρ' obtained from ρ by solely deleting one letter from A – this permutation contains all permutations in the former category while also being strictly contained in ρ. Similarly, we have

$$\mu(\tau, \rho) = - \sum_{z \in [\tau, \rho)} \mu(\tau, z) = \sum_{z \in [\tau, \rho']} \mu(\tau, z) = 0$$

□

An interesting line of inquiry would be to classify the intervals for which the Möbius function is 0 – both [5] and [6] answer these questions to some extent

0.5 Acknowledgements

I would like to thank Rik Sengupta for his invaluable help and guidance. In addition, I thank the MIT-PRIMES program and Dr. Khovanova for providing the project and providing many useful comments. I also thank Prof. Jacob Fox for helpful suggestions.

Bibliography

[1] H. Erikkson, K. Erikkson, S. Linusson, J. Wastlund. *Dense Packing of Patterns in a Permutation* (2007)

[2] A. Burstein, P. Hästö, T. Mansour. *Packing Patterns into Words* (2002) The Electronic Journal of Combinatorics Volume 9 Issue 2

[3] D. Kleitman, D. Kwiatkowski. *A lower bound on the length of a sequence containing all permutations as subsequences* (1976) Journal of Combinatorial Theory, Series A

[4] A. Miller *Asymptotic bounds for permutations containing many different patterns* Journal of Combinatorial Theory, Series A

[5] E. Steingrimsson, B. Tenner. *The Möbius Function of the Permutation Pattern Poset* (2009) http://arxiv.org/abs/0902.4011

[6] J. Smith. *On the Möbius Function of Permutations With One Descent* (2013) http://arxiv.org/abs/1306.5926