involve

a journal of mathematics

Note on superpatterns

Daniel Gray and Hua Wang

msp

2016 vol. 9, no. 5

INVOLVE 9:5 (2016)
dx.doi.org/10.2140/involve.2016.9.797

Note on superpatterns

Daniel Gray and Hua Wang

(Communicated by Joshua Cooper)

Given a set P of permutations, a P-superpattern is a permutation that contains every permutation in P as a pattern. The study of the minimum length of a superpattern has been of interest. For P being the set of all permutations of a given length, bounds on the minimum length have been improved over the years, and the minimum length is conjectured to be asymptotic with k^2/e^2. Similar questions have been considered for the set of layered permutations. We consider superpatterns with respect to packing colored permutations or multiple copies of permutations. Some simple but interesting observations will be presented. We also propose a few questions.

1. Introduction

Given a permutation π of length n, a pattern σ is said to be contained in π, or σ *occurs* in π, if a subsequence of π is order isomorphic to σ. For instance, the permutation $\pi = 51342$ contains two occurrences of the pattern $\sigma = 321$ as the subsequences 532 and 542. Much effort has been devoted to the study of occurrences of patterns in a permutation, most of which involves studying permutations which avoid a particular pattern, i.e., pattern avoidance.

As a symmetric problem to pattern avoidance, the concept of a superpattern concerns packing all patterns from a given set into a single permutation.

Definition. Let P be a set of permutations. A P-superpattern is a permutation that contains π for every $\pi \in P$.

Superpatterns were first introduced in [Arratia 1999]. The natural question immediately following this definition is to find the minimum length of a P-superpattern. When P is the set of all permutations of length k, this minimum length is denoted by $\mathrm{sp}(k)$ and has been vigorously studied. The trivial upper bound of k^2 was improved to $\frac{2}{3}k^2$ in [Eriksson et al. 2002], and was conjectured to be asymptotic with $\frac{1}{2}k^2$. Later, it was shown in [Miller 2009] that $\mathrm{sp}(k) \le \frac{1}{2}k(k+1)$ through

MSC2010: 05A05.
Keywords: superpatterns, colored permutations.
This work was partially supported by grants from the Simons Foundation (#245307).

the construction of a zigzag k-superword. More recently, bounds on the minimum lengths of superpatterns containing all layered permutations were considered in [Gray 2015]. Constructions similar to that in [Miller 2009] were also used to study superpatterns containing simple patterns of length k in [Gray $\geq$ 2016].

Definition. An m-colored permutation χ of length n is a permutation of length n in which each element is assigned one of m distinct colors.

For example, let $\chi = 3_1 2_1 5_1 1_1 4_2$ be a 2-colored permutation, where 3, 2, 5, and 1 have color 1, while 4 has color 2. Analogous to the case of noncolored patterns, the colored pattern $\phi = 2_1 1_1 3_2$ occurs in χ as the subsequences $3_1 1_1 4_2$, $2_1 1_1 4_2$, and $3_1 2_1 4_2$, but not $3_1 2_1 5_1$.

Colored permutations are of interest in the study of patterns and pattern avoidance [Savage and Wilf 2006]. Packing densities of colored permutations were also recently considered [Just and Wang 2016].

In this note we consider superpatterns of different sets of colored permutations. Some elementary, but interesting, observations will be presented. We also propose some questions from these studies.

2. Superpatterns containing all colored permutations

Let $S(k, m)$ denote the set of m-colored permutations of length k and

$$\mathrm{sp}(k, m) = \min\bigl\{|p| : p \text{ is an } S(k, m)\text{-superpattern}\bigr\}.$$

The following presents a simple connection between $\mathrm{sp}(k, m)$ and $\mathrm{sp}(k)$.

Theorem 2.1. *For any positive integers k and m, we have*

$$\mathrm{sp}(k, m) = m \cdot \mathrm{sp}(k).$$

Proof. Let p' be an $S(k, m)$-superpattern, and denote by p'_i the subsequence of p' of color i (for any $1 \leq i \leq m$). Then p'_i, without the color, is a superpattern containing all noncolored patterns of length k. Consequently $|p'_i| \geq \mathrm{sp}(k)$ for any i and

$$|p'| = \sum_{i=1}^{m} |p'_i| \geq m \cdot \mathrm{sp}(k).$$

On the other hand, let p be a permutation of length $\mathrm{sp}(k)$ that contains all noncolored patterns of length k. Construct an m-colored permutation p'' from p by replacing each $1 \leq j \leq \mathrm{sp}(k)$ in p by the sequence

$$s_j := [m(j-1)+1]_1 [m(j-1)+2]_2 \cdots [m(j-1)+m]_m.$$

Note that $|p''| = m \cdot |p| = m \cdot \mathrm{sp}(k)$. For any pattern $\pi \in S(k, m)$, the noncolored version is contained in p and the corresponding colored pattern can be found in p'' by choosing corresponding digits in s_j with the required color. Thus,

$$\mathrm{sp}(k, m) \leq |p''| = m \cdot \mathrm{sp}(k). \qquad \square$$

For example, $p = 132$ is a superpattern containing all patterns of length 2, and p is of length $\mathrm{sp}(2) = 3$. An $S(2, 3)$-superpattern p'' can be constructed as

$$1_1 2_2 3_3 7_1 8_2 9_3 4_1 5_2 6_3.$$

As an immediate consequence of Theorem 2.1, the established asymptotic bounds for $\mathrm{sp}(k)$ apply directly to $\mathrm{sp}(k, m)$. The trivial asymptotic lower bound k^2/e^2 for $\mathrm{sp}(k)$ follows from

$$\binom{\mathrm{sp}(k)}{k} \geq k!$$

and a standard application of Stirling's approximation for factorials [Arratia 1999].

Corollary 2.2. *For any positive integers k and m,*

$$mk^2/e^2 \leq \mathrm{sp}(k, m) \leq \tfrac{1}{2}mk(k+1).$$

Remark. The arguments in Theorem 2.1 establish the same relationship between the colored and noncolored versions of superpatterns containing any particular subset of the length-k permutations, such as the layered permutations [Gray 2015] and simple and alternating permutations [Gray $\geq$ 2016], and consequently provide bounds on the minimum lengths of these colored superpatterns.

3. Monochromatic and nonmonochromatic patterns

Let $NMS(k, m)$ be the set of nonmonochromatic m-colored patterns of length k and $MS(k, m)$ be the set of all monochromatic m-colored patterns of length k. Then, $S(k, m)$ is the disjoint union of $NMS(k, m)$ and $MS(k, m)$. It is easy to see that

$$|MS(k, m)| = mk!,$$

and consequently,

$$\begin{aligned} |NMS(k, m)| =& |S(k, m)| - |MS(k, m)| = m^k k! - mk! = (m^k - m)k! \\ =& (m^{k-1} - 1)\,|MS(k, m)|. \end{aligned}$$

Given any $NMS(k, m)$-superpattern of length n, we must have

$$\binom{n}{k} \geq (m^k - m)k!,$$

implying (by way of a standard application of Stirling's approximation for factorials)

$$n \geq mk^2/e^2,$$

the same asymptotic lower bound for $\mathrm{sp}(k, m)$ for general $S(k, m)$-superpatterns.

Letting

$$\mathrm{nmsp}(k, m) = \min\bigl\{|p| : p \text{ is an } NMS(k, m)\text{-superpattern}\bigr\},$$

we have the simple consequence that

$$mk^2/e^2 \leq \mathrm{nmsp}(k, m) \leq \mathrm{sp}(k, m) \leq \tfrac{1}{2}mk(k+1). \tag{1}$$

On the other hand, exactly the same argument as that of Theorem 2.1 implies

$$\operatorname{msp}(k, m) = m \cdot \operatorname{sp}(k), \tag{2}$$

where

$$\operatorname{msp}(k, m) = \min\bigl\{|p| : p \text{ is an } MS(k, m)\text{-superpattern}\bigr\}.$$

Remark. Equations (1) and (2) imply, in addition to the semitrivial bounds of $\operatorname{msp}(k, m)$ and $\operatorname{nmsp}(k, m)$, that

$$\operatorname{msp}(k, m) = m \cdot \operatorname{sp}(k) = \operatorname{sp}(k, m) \geq \operatorname{nmsp}(k, m), \tag{3}$$

a rather surprising fact given that $|NMS(k, m)| = (m^{k-1} - 1)\,|MS(k, m)|$.

While it may be a bit unexpected to see that $\operatorname{msp}(k, m) = \operatorname{sp}(k, m)$, a natural question follows.

Question 3.1. *Does strict inequality hold in* (3)?

In the special case for $k = 2$, the proposition below answers Question 3.1 in the affirmative.

Proposition 3.2. *For any positive integer* m, *we have*

$$3m = \operatorname{msp}(2, m) > 3m - 2 \geq \operatorname{nmsp}(2, m).$$

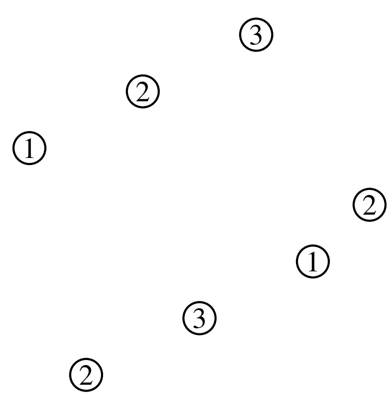
Proof. Clearly, $\operatorname{sp}(2) = 3$, and hence $\operatorname{msp}(2, m) = m \cdot \operatorname{sp}(2) = 3m$. Let p be a permutation of length $3m - 2$ defined as

$$[2m-1]_1 1_2 [2m]_2 2_3 \cdots [3m-3]_{m-1} [m-1]_m [3m-2]_m m_1 [m+1]_2 [m+2]_3 \cdots [2m-2]_{m-1}.$$

For instance, if $m = 3$, then

$$p = 5_1 1_2 6_2 2_3 7_3 3_1 4_2,$$

and the graph of p is depicted below:



In general, the graph of p will be the disjoint union of two increasing subsequences. The top row is of length m and the bottom row is of length $2m-2$, and every entry of the top row is larger than every entry of the bottom row. The entries of p alternate from the top row to the bottom row until there are m entries in the top row. Then, $m-2$ more entries are added to the bottom row. The i-th entry of the top row will have color i for $1 \le i \le m$, while the i-th entry of the bottom row will have color $i+1 \pmod m$.

For $i, j \in [1, m]$ with $i \neq j$, the pattern $1_i 2_j$ is contained in the bottom row, the only exception being the pattern $1_1 2_j$, which is contained in the top row. The pattern $2_i 1_j$ can be found by selecting the unique entry colored i from the top row, and taking an entry in the bottom row colored j which lies to the right of the entry just selected. Then, p is an $NMS(2, m)$-superpattern. □

To answer Question 3.1 in general appears to be very difficult. In an effort to further understand the relationship between monochromatic and nonmonochromatic superpatterns, we also point out the following.

Proposition 3.3. *For any positive integers $k \ge 2$ and m, we have*

$$\mathrm{msp}(k-1, m) \le \mathrm{nmsp}(k, m) \le \mathrm{msp}(k, m).$$

Proof. The second inequality is implied by the remark on page 800. To see the first inequality, let q be an m-colored pattern of length k whose first $k-1$ entries are colored by color i and whose k-th entry is colored by $j \neq i$. Then, the first $k-1$ entries of q form a monochromatic pattern of length $k-1$.

Since q is a nonmonochromatic m-colored pattern of length k, it must be contained in any $NMS(k, m)$-superpattern. Noting that we could have colored the first $k-1$ entries of q monochromatically using any of the m colors, any $NMS(k, m)$-superpattern must contain all monochromatic m-colored patterns of length $k-1$. Thus, $\mathrm{nmsp}(k, m) \ge \mathrm{msp}(k-1, m)$. □

4. Packing multiple copies of all patterns

The idea of superpatterns lies in the fact that they contain each permutation (from a given set of permutations) at least once. A natural generalization seems to be superpatterns that contain each permutation at least a given number of times.

Definition. For a given set P of permutations, a P_ℓ-superpattern is a permutation containing each pattern $\pi \in P$ at least ℓ times.

Define $\mathrm{sp}_\ell(k)$, $\mathrm{sp}_\ell(k, m)$, $\mathrm{msp}_\ell(k, m)$, and $\mathrm{nmsp}_\ell(k, m)$ accordingly. Some trivial facts follow immediately:

- $\mathrm{sp}_\ell(k, m) = m \cdot \mathrm{sp}_\ell(k)$. This can be seen by following exactly the same argument as that of Theorem 2.1.

- $\mathrm{sp}_\ell(k) \leq \ell \cdot \mathrm{sp}(k)$ and $\mathrm{sp}_\ell(k,m) \leq \ell \cdot \mathrm{sp}(k,m)$. For permutations p of length n and q of length m, the direct sum $p \oplus q$ is the permutation that has the first n entries from p and the next m entries from entries of q shifted by n. That is,

$$p \oplus q = p_1 p_2 \cdots p_n (q_1+n)(q_2+n)\cdots(q_m+n).$$

  Given a permutation p of length $\mathrm{sp}(k)$ that contains all patterns of length k, the permutation $\bigoplus_{i=1}^{\ell} p$ clearly contains each length-k pattern in each of the ℓ summands. Hence, $\mathrm{sp}_\ell(k) \leq \ell \cdot \mathrm{sp}(k)$. A similar argument holds for $\mathrm{sp}_\ell(k,m) \leq \ell \cdot \mathrm{sp}(k,m)$.

- $\mathrm{sp}_m(k) \leq \mathrm{msp}(k,m)$. This can be seen from removing the colors of an $MS(k,m)$-superpattern.

The asymptotic lower bounds, for k large and ℓ constant, of $\mathrm{sp}_\ell(k)$ or $\mathrm{sp}_\ell(k,m)$ stay the same as $\mathrm{sp}(k)$ or $\mathrm{sp}(k,m)$. Given that the multiple copies of patterns need not be disjoint, it is natural to ask for improvement of the upper bounds above. The existing constructions (that provided upper bounds for the minimum lengths of various superpatterns) such as those in [Arratia 1999; Gray 2015; ≥ 2016] do not directly generalize to the case of packing multiple copies of every permutation. We conclude this note by showing a nontrivial upper bound for the sp_ℓ function.

Definition. For $k,n \in \mathbb{N}$, let $q = q_1q_2q_3\cdots q_k$ be a pattern of length k and let $w = w_1w_2w_3\cdots w_n$ be a word of length n. We say that q is "contained exactly in w" if there is a subsequence of length k, say

$$(w_{i_1}, w_{i_2}, \ldots, w_{i_k}),$$

such that $w_{i_j} = q_j$ for all $1 \leq j \leq k$.

Theorem 4.1. *For $k,\ell \in \mathbb{N}$, we have*

$$\mathrm{sp}_\ell(k) \leq \begin{cases} \frac{1}{2}(k+1)(k+\ell-1) & \text{if } k \text{ is odd},\\ \frac{1}{2}(k+1)(k+\ell-1) & \text{if } k \text{ is even and } \ell \text{ is odd},\\ \frac{1}{2}(k+1)(k+\ell-1)+1 & \text{if } k \text{ is even and } \ell \text{ is even}.\end{cases}$$

Proof. We begin with Allison Miller's construction [2009] of the zigzag k-superword. Let k_o (resp. k_e) be the smallest odd (resp. even) integer at least as large as k. We make the following definitions:

$$\bar{k}_o = 1357\cdots k_o \quad \text{and} \quad \bar{k}_e = k_e \cdots 8642.$$

Define

$$w = \bar{k}_o\bar{k}_e\bar{k}_o\bar{k}_e \cdots \bar{k}_o\bar{k}_e$$

if k is even or

$$w = \bar{k}_o\bar{k}_e\bar{k}_o\bar{k}_e \cdots \bar{k}_o\bar{k}_e\bar{k}_o$$

if k is odd, where the combined number of copies of $\bar{k}_o$ and $\bar{k}_e$ is exactly k. The word w is what Miller calls the zigzag k-superword.

Let q be a pattern of length k, and let $q+1$ be the permutation of the set $\{2, 3, 4, \ldots, k+1\}$ obtained by adding 1 to each entry of q. Number the runs of w from left to right in increasing order and let $m(q)$ be the number of runs needed (in $\bar{k}_o\bar{k}_e\bar{k}_o\bar{k}_e\cdots$) to contain q. Miller shows that

$$m(q) + m(q+1) \leq 2k+1, \tag{4}$$

but in fact, the same steps can be used to show equality in (4). Hence, either $m(q)$ or $m(q+1)$ is at most k, which implies either q or $q+1$ is contained exactly in w.

Now, consider the finite word

$$w(\ell) = \begin{cases} \bar{k}_o\bar{k}_e\bar{k}_o\bar{k}_e\cdots\bar{k}_o\bar{k}_e & \text{if } k \text{ is even,} \\ \bar{k}_o\bar{k}_e\bar{k}_o\bar{k}_e\cdots\bar{k}_o\bar{k}_e\bar{k}_o & \text{if } k \text{ is odd,} \end{cases}$$

where the combined number of copies of $\bar{k}_o$ and $\bar{k}_e$ is exactly $k+\ell-1$. Suppose without loss of generality that $m(q) \leq k$, and recall, $m(q+1) = 2k+1-m(q)$. Since q is contained in the first $m(q)$ copies of $w(\ell)$, and each run of $w(\ell)$ is repeated every two times, there is another copy of q contained between the third run and the $(m(q)+2)$-th run, yet another copy of q contained between the fifth and $(m(q)+4)$-th runs, and so on for as long as we do not exceed $k+\ell-1$ runs. Thus, there are at least

$$1 + \left\lfloor \tfrac{1}{2}\big((k+\ell-1) - m(q)\big)\right\rfloor = \left\lfloor \tfrac{1}{2}\big(\ell + (k-m(q)) + 1\big)\right\rfloor$$

copies of q contained exactly in $w(\ell)$. Hence, if $k - m(q) \geq \ell - 1$, we successfully have at least ℓ copies of q. Then, let us suppose that $k - m(q) < \ell - 1$. For the same reason as above, there are at least

$$1 + \left\lfloor \tfrac{1}{2}\big((k+\ell-1) - (2k+1-m(q))\big)\right\rfloor = \left\lfloor \tfrac{1}{2}\big(\ell - (k-m(q))\big)\right\rfloor \geq 1$$

copies of $q+1$ contained exactly in $w(\ell)$. Hence, the combined number of copies of $m(q)$ and $m(q+1)$ is at least

$$\left\lfloor \tfrac{1}{2}\big(\ell + (k-m(q)) + 1\big)\right\rfloor + \left\lfloor \tfrac{1}{2}\big(\ell - (k-m(q))\big)\right\rfloor = \ell.$$

Finding a permutation p representing $w(\ell)$ is routine. Note that p will contain at least ℓ copies of q. Let us consider the length of $w(\ell)$. First suppose k is odd. Then, there are $\frac{1}{2}(k+1)$ entries each in $\bar{k}_o$ and $\bar{k}_e$. Miller shows that w is of length $\frac{1}{2}k(k+1)$, to which we add $\ell-1$ more runs. Hence, the length of $w(\ell)$ is

$$\tfrac{1}{2}k(k+1) + (\ell-1)\tfrac{1}{2}(k+1) = \tfrac{1}{2}(k+1)(k+\ell-1).$$

Now suppose that k is even. Then, there are $\frac{1}{2}k$ entries in $\bar{k}_e$ and $1+\frac{1}{2}k = \frac{1}{2}(k+2)$ entries in $\bar{k}_o$. If ℓ is odd, then we have added $\frac{1}{2}(\ell-1)$ copies each of $\bar{k}_e$ and $\bar{k}_o$.

Thus, the length of $w(\ell)$ is

$$\tfrac{1}{2}k(k+1)+\tfrac{1}{2}(\ell-1)\tfrac{1}{2}k+\tfrac{1}{2}(\ell-1)\tfrac{1}{2}(k+2)=\tfrac{1}{2}(k+1)(k+\ell-1).$$

If ℓ is even, then we have added $\frac{1}{2}(\ell-2)$ copies of $\bar{k}_e$ and $\frac{1}{2}\ell$ copies of $\bar{k}_o$. Therefore, the length of $w(\ell)$ is

$$\tfrac{1}{2}k(k+1)+\tfrac{1}{2}(\ell-2)\tfrac{1}{2}k+\tfrac{1}{2}\ell\tfrac{1}{2}(k+2)=\tfrac{1}{2}(k+1)(k+\ell-1)+1. \qquad \square$$

Remark. The above argument can be easily modified, by using the construction in Theorem 2.1, to provide less trivial upper bounds for $\mathrm{sp}_\ell(k,m)$.

Remark. It is also interesting to note that, if one takes a superpattern from $S(k,m)$ achieving $\mathrm{sp}(k,m)$ and removes colors, the resulting noncolored permutation is a superpattern that contains each k-pattern m^k times (since there are m^k different ways to color a k-pattern with m colors).

Acknowledgement

We greatly appreciate many helpful suggestions and corrections from the anonymous referee.

References

[Arratia 1999] R. Arratia, "On the Stanley–Wilf conjecture for the number of permutations avoiding a given pattern", *Electron. J. Combin.* **6** (1999), Note N1. MR 1710623 Zbl 0922.05002

[Eriksson et al. 2002] H. Eriksson, K. Eriksson, S. Linusson, and J. Wästlund, "Dense packing of patterns in a permutation", in *FPSAC '02: 14th Conference on Formal Power Series and Algebraic Combinatorics* (Melbourne 2002), 2002.

[Gray 2015] D. Gray, "Bounds on superpatterns containing all layered permutations", *Graphs Combin.* **31**:4 (2015), 941–952. MR 3357666 Zbl 1316.05001

[Gray ≥ 2016] D. Gray, "Simple permutations, alternating permutations, and the superpatterns that contain them", preprint. To appear in *Ars Combin.*

[Just and Wang 2016] M. Just and H. Wang, "Note on packing patterns in colored permutations", *Online J. Anal. Comb.* **11** (2016). Zbl 06574956

[Miller 2009] A. Miller, "Asymptotic bounds for permutations containing many different patterns", *J. Combin. Theory Ser. A* **116**:1 (2009), 92–108. MR 2469250 Zbl 1177.05012

[Savage and Wilf 2006] C. D. Savage and H. S. Wilf, "Pattern avoidance in compositions and multiset permutations", *Adv. in Appl. Math.* **36**:2 (2006), 194–201. MR 2199988 Zbl 1087.05002

Received: 2015-05-01 Revised: 2015-09-10 Accepted: 2015-09-17

dgray1@ufl.edu *Department of Mathematics, University of Florida, Gainesville, FL 32611, United States*

hwang@georgiasouthern.edu *Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA 30460, United States*

involve

msp.org/involve

INVOLVE YOUR STUDENTS IN RESEARCH

Involve showcases and encourages high-quality mathematical research involving students from all academic levels. The editorial board consists of mathematical scientists committed to nurturing student participation in research. Bridging the gap between the extremes of purely undergraduate research journals and mainstream research journals, *Involve* provides a venue to mathematicians wishing to encourage the creative involvement of students.

See inside back cover or msp.org/involve for submission instructions. The subscription price for 2016 is US $160/year for the electronic version, and $215/year (+$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Involve (ISSN 1944-4184 electronic, 1944-4176 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

Involve peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

involve

2016 vol. 9 no. 5

1944-4176(2016)9:5;1-0