Advances in Applied Mathematics 81 (2016) 202–211

Contents lists available at ScienceDirect

Advances in Applied Mathematics

www.elsevier.com/locate/yaama

Some results on superpatterns for preferential arrangements

Yonah Biers-Ariel [a], Yiguang Zhang [b], Anant Godbole [c,*]

[a] *Department of Mathematics, Rutgers University, United States*
[b] *Department of Statistics, Stanford University, United States*
[c] *Department of Mathematics & Statistics, East Tennessee State University, United States*

A R T I C L E I N F O

Article history:
Received 29 February 2016
Received in revised form 26 July 2016
Accepted 23 August 2016
Available online 31 August 2016

MSC:
05D99
05D40
05A05

Keywords:
Superpattern
Complete words
Preferential arrangement
Permutation

A B S T R A C T

A *superpattern* is a string of characters of length n over $[k] = \{1, 2, \ldots, k\}$ that contains as a subsequence, and in a sense that depends on the context, all the smaller strings of length k in a certain class. We prove structural and probabilistic results on superpatterns for *preferential arrangements*, including (i) a theorem that demonstrates that a string is a superpattern for all preferential arrangements if and only if it is a superpattern for all permutations; and (ii) a result that is reminiscent of a still unresolved conjecture of Alon on the smallest permutation on $[n]$ that contains all k-permutations with high probability.

1. Introduction and statement of results

A *superpattern* is a string of characters of length n that contains as a subsequence, and in a sense that depends on the context, all the smaller strings of length k in a certain

* Corresponding author.
E-mail address: godbolea@etsu.edu (A. Godbole).

http://dx.doi.org/10.1016/j.aam.2016.08.004
0196-8858/

class. Specifically, given a set X and a class $\mathcal{R}$ such that each object in $\mathcal{R}$ is a string of k elements in X, a *superpattern* is a string that contains all $p \in \mathcal{R}$ as subsequences. For example, with $X = \{1, 2\}$ and $\mathcal{R} = \{11, 12, 21, 22\}$,

$$1221$$

is a superpattern.

In this paper, we present some results on superpatterns for *preferential arrangements*, or word-patterns. Key references in this area are [3,5,8]. Preferential arrangements (p.a.'s) of length k over $X = [d] := \{1, 2, \ldots, d\}$ are k-strings with entries from $[d]$, for which order isomorphic representations are considered to be equivalent. For example if $k = 3$, $d = 2$, there are seven preferential arrangements, *viz.* 111, 112, 121, 211, 122, 212, and 221. If $k = d = 3$, the thirteen preferential arrangements (enumerated whenever $d = k$ by the ordered Bell numbers) are 112, 121, 211, 122, 212, 221, 111, and the six permutations 123, 132, 213, 231, 312, and 321. Note that, for example, the strings 112, 113, and 223 are order isomorphic, so above we just list the preferential arrangement 112, expressed in the traditional lexicographically minimal fashion, also known as a dense ranking system. If $k = 3$, $d \geq 4$, there are still only 13 p.a.'s, since, for example, with $k = 3$ and $d = 4$, the six strings 112, 113, 114, 223, 224, 334 are each equivalent to the p.a. 112. To see this fact in general, if we have more characters than the length of the word, any such word omits characters and we can thus represent it with an isomorphic word using characters from $\{1, \ldots, k\}$, and thus the number of preferential arrangements for $k < d$ is equal to the arrangements for $k = d$.

A superpattern for preferential arrangements of length k over $[d]$ is an n-long string over the alphabet $[d]$, that contains, as a subsequence, each of the preferential arrangements of length k over $[d]$ in any one of its order isomorphic forms. For example, a string such as 3213213 is a superpattern for $k = d = 3$ *or with* $k = 3$; $d = 4$, and 1231241 is a superpattern with, e.g., $k = 3$; $d = 5$ or $k = 3$; $d = 4$. Let $n(k, d)$ be the length of the shortest superpattern for all p.a.'s of length k over $[d]$.

Now, let us define what we consider to be another natural object: Let $\nu(k, d)$ be the length of shortest superpattern for k-long p.a.'s *when each of the letters in* $[d]$ *must be used at least once in the superpattern.* For $k = 3$, the examples 1231231, 1231241, and 2353134, as well as the fact that the p.a. 111 can never occur with $n = 7$; $d \geq 6$ show that $\nu(3, d) = 7$ for $d = 3, 4, 5$ and that $\nu(3, d) = 7 + (d - 5)$ for $d \geq 6$. (Technically, we have just shown that $\nu(3, 4) \leq 7$; $\nu(3, 5) \leq 7$. A proof that these values equal 7 is not too difficult, and is omitted.) For $k = 4$ the situation is more complex: We have $\nu(4, 4) = 12$, as seen in Section 2, but $\nu(4, 6) \leq 11$, as seen via the example 43514342634. This example also shows that the assertion in [3] that $n(k, d) = n(k, k)$ for $d > k$ needs further qualification.

Open question: For fixed small values of k, d, calculate $\nu(k, d)$. Establish upper and lower bounds on $\nu(k, d)$.