Theoretical Computer Science ••• (••••) •••–•••

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Unary patterns under permutations

James Currie [a], Florin Manea [b,*], Dirk Nowotka [b], Kamellia Reshadi [b]

[a] Department of Mathematics, University of Winnipeg, Canada
[b] Kiel University, Department of Computer Science, Kiel, Germany

ARTICLE INFO

Article history:
Received 3 January 2018
Received in revised form 14 May 2018
Accepted 25 May 2018
Available online xxxx

Keywords:
Combinatorics on words
Avoidable patterns
Patterns under permutations

ABSTRACT

Thue characterized completely the avoidability of unary patterns. Adding function variables gives a general setting capturing avoidance of powers, avoidance of patterns with palindromes, avoidance of powers under coding, and other questions of recent interest. Unary patterns with permutations have been previously analysed only for lengths up to 3. Consider a pattern $p = \pi_{i_1}(x)\dots\pi_{i_r}(x)$, with $r \geq 4$, x a word variable over an alphabet Σ and π_{i_j} function variables, to be replaced by morphic or antimorphic permutations of Σ. If $|\Sigma| \geq 3$, we show the existence of an infinite word avoiding all pattern instances having $|x| \geq 2$. If $|\Sigma| = 3$ and all π_{i_j} are powers of a single morphic or antimorphic π, the length restriction is removed. For the case when π is morphic, the length dependency can be removed also for $|\Sigma| = 4$, but not for $|\Sigma| = 5$, as the pattern $x\pi^2(x)\pi^{56}(x)\pi^{33}(x)$ becomes unavoidable. Thus, in general, the restriction on x cannot be removed, even for powers of morphic permutations. Moreover, we show that for every positive integer n there exists N and a pattern $\pi^{i_1}(x)\dots\pi^{i_n}(x)$ which is unavoidable over all alphabets Σ with at least N letters and π morphic or antimorphic permutation.

1. Introduction

The avoidability of patterns by infinite words is a core topic in combinatorics on words, going back to Thue [2,3]. Important results are surveyed in, e.g., [4,5].

Recently, a natural generalisation of classical patterns, in which functional dependencies between variables are allowed, has been considered [6–8]. More precisely, patterns consist of word variables, as usual, together with function variables (standing for either morphic or antimorphic extensions of permutations on the alphabet) which act on the words. For example, consider the pattern $x\pi(x)x\pi(x)$ whose instances are words $uvuv$ that consist of four parts of equal length, that is, $|u| = |v|$, where v is the image of u under some permutation of the alphabet. For example, $aab|bba|aab|bba$ (respectively, $aab|abb|aab|abb$) is an instance of $x\pi(x)x\pi(x)$ for the morphic (respectively, antimorphic) extension of permutation $a \mapsto b$ and $b \mapsto a$.

We note that, while patterns x^k describe all repetitions of some exponent k, patterns of the type $\pi^{i_1}(x)\dots\pi^{i_k}(x)$ describe words that have an intrinsic repetitive structure, hidden by the application of the different iterations of the function π, which encode of the original root of the repetition.

* Corresponding author.
E-mail addresses: j.currie@uwinnipeg.ca (J. Currie), flm@informatik.uni-kiel.de (F. Manea), dn@informatik.uni-kiel.de (D. Nowotka), kre@informatik.uni-kiel.de (K. Reshadi).

https://doi.org/10.1016/j.tcs.2018.05.033

Patterns with involutions were studied in [6,7]; motivation for considering involutions includes word reversal and DNA/RNA complementation. The main result obtained was that for each unary pattern with one variable involution, one can identify all alphabets over which it is avoidable. In the more general setting of patterns with permutations, the only results obtained so far (see [8]) regarded cube-like patterns under morphisms or antimorphisms (anti-/morphisms, for short) which are powers of a single (variable) permutation, i.e., patterns of the form $\pi^i(x)\pi^j(x)\pi^k(x)$, where $i, j, k \geq 0$. The avoidability of such patterns was completely characterised: for each $\pi^i(x)\pi^j(x)\pi^k(x)$ one can determine exactly the alphabets over which the pattern is avoidable. Contrary to both the classical and to the involution settings, where once a pattern is avoidable for some alphabet size it remains avoidable in larger alphabets, a cubic pattern with permutations may become unavoidable over a larger alphabet.

1.1. Our contribution

We extend the results of [8] as follows.

First, we construct a ternary word that avoids all patterns $\pi_{i_1}(x)\ldots\pi_{i_r}(x)$ where $r \geq 4$, x a word variable over some alphabet Σ, with $|x| \geq 2$ and $|\Sigma| \geq 3$, and the π_{i_j} function variables that may be replaced by anti-/morphic permutations of Σ. This is the first result where the avoidability of patterns involving more functions, which are not powers of the same initial variable permutation, has been shown; even more, we do not restrict these functions so that all have the same type: we can mix both morphic and antimorphic permutations.

On the down side, the result above only works when we restrict the length of x to be at least 2. However, we also show that such a restriction is needed. Indeed, for each $n \geq 1$ there exists a unary pattern $\pi_1(x)\ldots\pi_n(x)$ where all functions are powers of the same anti-/morphic permutation π, i.e., $\pi_j = \pi^{i_j}$ with $1 \leq j \leq n$, and an integer N such that $\pi^{i_1}(x)\ldots\pi^{i_n}(x)$ has as instances all the words of length n over an alphabet of size at least N; in other words, $\pi^{i_1}(x)\ldots\pi^{i_n}(x)$ is unavoidable over all alphabets Σ with $|\Sigma| \geq N$.

In between these two results, we show that all patterns $\pi^{i_1}(x)\ldots\pi^{i_n}(x)$ with $n \geq 4$ under anti-/morphic permutations are avoidable in Σ_3. Similarly, all patterns $\pi^{i_1}(x)\ldots\pi^{i_n}(x)$ under morphic permutations are also avoidable in Σ_4, but not in Σ_5. So, just like in the case of cubes with permutations, there are patterns under anti-/morphic permutations (including the eventually unavoidable patterns we construct) which are avoidable in small alphabets (e.g., in Σ_3 or Σ_4) but become unavoidable in larger alphabets. On the other hand, unlike the case of cubes with permutations, where there exist patterns unavoidable in Σ_2 and Σ_3 (e.g., $x\pi(x)\pi^2(x)$, see [8]), all unary patterns of length at least 4 under anti-/morphic permutations are avoidable in both Σ_2 (see [7]) and Σ_3, and, in the case of morphic permutations, in Σ_4 as well. Note that 4 is the largest integer i such that all patterns of length 4 under morphic permutations are avoidable in Σ_i.

2. Definitions

We freely use the usual notations of combinatorics on words (see, for instance, [4]). Define alphabets $\Sigma_k = \{0, \ldots, k-1\}$ and $\Sigma'_k = \{1, 2, \ldots, k\}$. We use w^R, to denote the reversal of word w.

A morphism f (respectively, antimorphism) of Σ_k^* is defined by its values on letters; $f(uv) = f(u)f(v)$ (respectively, $f(uv) = f(v)f(u)$) for all words $u, v \in \Sigma_k^*$. When we define an anti-/morphism it is enough to define $f(a)$, for all $a \in \Sigma_k$. If the restriction of f to Σ_k, is a permutation of Σ_k, we call f an anti-/morphic permutation. Denote by $\mathbf{ord}(f)$ the order of f, i.e., the minimum positive integer m such that f^m is the identity. If $\mathbf{ord}(f) = 2$, we call f an involution. If $a \in \Sigma_k$ is a letter, the order of a with respect to f, denoted $\mathbf{ord}_f(a)$, is the minimum number m such that $f^m(a) = a$.

A pattern which involves functional dependencies is a term over (word) variables and function variables (where concatenation is an implicit functional constant); a pattern with only one word variable is called unary. For example, $x\pi(x)\pi(\pi(x))x = x\pi(x)\pi^2(x)x$ is a unary pattern involving the variable x and the function variable π. An instance of a pattern p in Σ_k is the result of substituting every variable by a word in Σ_k^+ and every function variable by a function over Σ_k^*. A pattern is avoidable in Σ_k if there is an infinite word over Σ_k that does not contain any instance of the pattern.

In this paper, we consider patterns with morphic and antimorphic permutations, that is, all function variables are substituted by morphic or antimorphic permutations only.

The infinite Thue–Morse word $\mathbf{t}$ (see [2]) is defined as $\mathbf{t} = \lim_{n\to\infty} \phi_t^n(0)$, for the morphism $\phi_t : \Sigma_2^* \to \Sigma_2^*$ where $\phi_t(0) = 01$ and $\phi_t(1) = 10$. It is well-known (see, for instance, [4]) that $\mathbf{t}$ avoids the patterns xxx (cubes) and $xyxyx$ (overlaps).

Let $\mathbf{h}$ be the infinite word defined as $\mathbf{h} = \lim_{n\to\infty} \phi_h^n(0)$, where $\phi_h : \Sigma_3^* \to \Sigma_3^*$ is a morphism due to Thue [2], which was rediscovered and studied also by Hall [9], defined by $\phi_h(0) = 012$, $\phi_h(1) = 02$ and $\phi_h(2) = 1$. For the simplicity of the exposure, if $\mathbf{h} = \prod_{i=0}^{\infty} h_i$ with $h_i \in \Sigma_3$, we define the infinite word $\mathbf{v}$ over Σ'_3 as $\mathbf{v} = \prod_{i=0}^{\infty} v_i$, with $v_i = h_i + 1$. The infinite word $\mathbf{v}$ (respectively, the word $\mathbf{h}$) avoids squares xx and does not contain the factors 121 and 323 (respectively, the factors 010 and 212).

We investigate the factors of an infinite word $\mathbf{g}$ that have the form

$$\pi_{i_1}(x)\pi_{i_2}(x)\ldots\pi_{i_r}(x)$$

Please cite this article in press as: J. Currie et al., Unary patterns under permutations, Theoret. Comput. Sci. (2018), https://doi.org/10.1016/j.tcs.2018.05.033