More on square-free words obtained from prefixes by permutations

Pascal Ochem
LIRMM, CNRS, Univ. Montpellier 2
ochem@lirmm.fr

May 23, 2013

Abstract

An infinite square-free word w over the alphabet $\Sigma_3 = \{0, 1, 2\}$ is said to have a k-stem σ if $|\sigma| = k$ and $w = \sigma w_1 w_2 \cdots$ where for each i, there exists a permutation π_i of Σ_3 which extended to a morphism gives $w_i = \pi_i(\sigma)$. Harju proved that there exists an infinite k-stem word for $k = 1, 2, 3, 9$ and $13 \leq k \leq 19$, but not for $4 \leq k \leq 8$ and $10 \leq k \leq 12$. He asked whether k-stem words exist for each $k \geq 20$. We give a positive answer to this question. Currie has found another construction that answers Harju's question.

1 Introduction

An infinite square-free word w over the alphabet $\Sigma_3 = \{0, 1, 2\}$ is said to have a k-stem σ if $|\sigma| = k$ and $w = \sigma w_1 w_2 \cdots$ where for each i, there exists a permutation π_i of Σ_3 which extended to a morphism gives $w_i = \pi_i(\sigma)$. Harju [3] proved that there exists an infinite k-stem word for $k = 1, 2, 3, 9$ and $13 \leq k \leq 19$, but not for $4 \leq k \leq 8$ and $10 \leq k \leq 12$ and asked whether k-stem words exist for each $k \geq 20$. We construct k-stem words for each $20 \leq k \leq 10000$ in Section 3 and for every $k \geq 23$ in Section 4. Currie [2] has found another construction that answers Harju's question.

Let $t = 012021012102012021020121012\ldots$ denote the fixed point of the morphism $0 \mapsto 012$, $1 \mapsto 02$, $2 \mapsto 1$. By definition, t contains neither 010 nor 212 as a factor. Harju [3] also asked whether t has a k-stem factorization for some $k \geq 3$. We give a negative answer in Section 2. This result has also been obtained by Harju and Müller [4].

2 k-stem factorization of t

Theorem 1 *No suffix of t admits a k-stem factorization for any $k \geq 3$.*

Proof. By previous results [3], we only need to consider the cases $k = 9$ and $k \geq 13$.

A computer check shows that no factor f of t of length 18 is such that the suffix of length 9 of f is a permutation of the prefix of length 9 of f. This rules out the case $k = 9$.

A computer check shows that every factor f of t of length 12 contains a factor $a0a$ with $a \in \Sigma_3$. By symmetry, it also contains a factor $b2b$ with $b \in \Sigma_3$. Remember that 010 and 212 are not factors of t. A permutation of Σ_3 mapping 0 to 1 (resp. mapping 2 to 1) cannot be applied to f, since it would produce a factor $c1c$ with $c \in \Sigma_3$ that cannot appear in t . There remain two possible permutations, namely the identity and the permutation swapping 0 and 2, but an infinite square free word cannot be obtained by a concatenation of only two distinct factors. This rules out the case $k \geq 13$.

3 k-stem words for $20 \leq k \leq 10000$

Theorem 2 *There exist k-stem words for every $20 \leq k \leq 10000$.*

Proof. Let π be the permutation (012). We say that a morphism $h : \Sigma_3^* \to \Sigma_3^*$ is circular if $h(1) = \pi(h(0))$ and $h(2) = \pi(h(1))$. For every $20 \leq k \leq 10000$, we found a word w_k such that $|w_k| = c_k \times k$ and the circular morphism m defined by $m(0) = w_k$ is square-free. We have $c_k = 8$ for $20 \leq k \leq 22$ and $c_k = 1$ for $23 \leq k \leq 10000$. Square-freeness is checked using the result of Crochemore [1] that a uniform morphism h is square-free if and only if the h-images of square-free words of length 3 are square-free. Since we consider circular morphisms, we only need to check the images of 010 and 012.

These are our words w_k for $20 \leq k \leq 22$, where $|w_k| = 8k$.

$w_{20} =$ 012102010210121021201020121012010201202120102120210201021012
0212010201202120121012021012102120210201210120212012101201 02
120210121021202102010212010201202120 1210

$w_{21} =$ 0120210201021201020120210120102012102010212012102120210120 21201
0210120102012102010210120210201021201020121021201210120210 12102
120102101210201210120210201202120102120210

$w_{22} =$ 0120210201021202102012021012010201210120102120121021202101210 21201
0210120102012101201021012021020102120210201210212012101202120 12102
1201021012102010210120210201202120102012 0210

Consider now the case $k \geq 23$, where $|w_k| = k$. Let $t' = 012021020121012\ldots$ denote the infinite suffix of t obtained from t by deleting the first 12 letters. To speed up the search of a suitable w_k, we impose that $w_k = pr120210$ where p is the prefix of length $k - 22$ of t' and r belongs to the set S of size 13 below, except that $r = 2102010210121020$ for $k = 26$.

$S = \{0102120121020102, 0120102120121020, 0212012101201020,$
$1012010212012102, 1021201021012102, 1201020121020102, 1202120121020102,$

$1210120212012102, 1210201202120102, 1210201210120102, 2010210121020102,$
$2102120121020102, 2120102101201020\}$

4 k-stem words for large k

Theorem 3 *There exist k-stem words for every $k \geq 1$ except for $4 \leq k \leq 8$ and $10 \leq k \leq 12$*

Proof. Consider the following morphism d, having two possible images for each letter: one image of length 17 and one image of length 18.

$$0 \mapsto \begin{cases} 01202120102120210 \\ 012021020102120210 \end{cases}$$

$$1 \mapsto \begin{cases} 12010201210201021 \\ 120102101210201021 \end{cases}$$

$$2 \mapsto \begin{cases} 20121012021012102 \\ 201210212021012102 \end{cases}$$

Again, using the result of Crochemore [1], d is shown to be square-free by checking that the d-images of square-free words of length $\max\left(3, \left\lceil \frac{18-3}{17} \right\rceil\right) = 3$ are square-free. Since the restriction of d to images of length 17 (resp. 18) is circular, we only need to check the images of 010 and 012 are square-free. For each of the factors 010 and 012, we actually have 2^3 images to check since each of the letters can be mapped either to its image of length 17 or 18.

If m is a square-free circular morphism, then for every d-image w_0 of $m(0)$, the circular morphism defined by $0 \mapsto w_0$ is square-free. This means that given a k-uniform square-free circular morphism, we can construct a k'-uniform square-free circular morphism for every k' such that $17k \leq k' \leq 18k$.

Now we prove that there exist k-uniform square-free circular morphisms for every $k \geq 23$. We start with the cases $k \in [23, 10000]$ which are proved in the previous section. They imply the cases $k \in \bigcup_{23 \leq p \leq 10000} [17p, 18p]$, i.e., $k \in [391, 180000]$. We then obtain every $k \geq 23$ by induction.

5 Concluding remarks

We have proved that there exist infinite square-free ternary words with a k-stem factorization for every k except $4 \leq k \leq 8$ and $10 \leq k \leq 12$. We conjecture that there exist k-stem words of the form $w_k = pr120210$ described in the proof of Theorem 2 for every $k \geq 23$, rather than $23 \leq k \leq 10000$. Before we found the morphism of the proof of Theorem 3, we pushed the verification to up to 10000 in order to find a way to prove this conjecture, but the proof of Theorem 3 only requires a verification for $23 \leq k \leq 390$.

From the proof of Theorem 3, we see that the number of k-uniform square-free circular morphisms is exponential in k, at least about $\binom{2k/35}{k/35} \approx 2^{2k/35}$. We conjecture the following:

Conjecture 4 *The growth rate of ternary words defining a square-free circular morphism exists and is equal to the growth rate* $1.3017\ldots$ *of ternary square-free words.*

See Shur [5] for more information on the growth rate of ternary square-free words.

References

[1] M. Crochemore. Sharp characterizations of squarefree morphisms, *Theoret. Comput. Sci.* **18(2)** (1982), 221–226.

[2] J. Currie. Infinite ternary square-free words concatenated from permutations of a single word, *Theoret. Comput. Sci.* **482** (2013), 1–8.

[3] T. Harju. Square-free words obtained from prefixes by permutations, *Theoret. Comput. Sci.* **429** (2012), 128–133.

[4] T. Harju and M. Müller. Square-free words generated by applying permutations to a prefix, in Proceedings of the Second Russian Finnish Symposium on Discrete Mathematics, RuFiDim II, (V. Halava, J. Karhumäki, Y. Matiyasevich, eds.) (2012), 86–91.

[5] A. Shur. Growth rates of complexity of power-free languages, *Theoret. Comput. Sci.* **411(34-36)** (2010), 3209–3223.