Theoretical Computer Science 38 (1985) 117-122
North-Holland

NOTE

A CHARACTERIZATION OF POWER-FREE MORPHISMS

Michel LECONTE
Université PARIS VII (LITP), 2, Place Jussieu, 75221 Paris Cedex 05, France

Communicated by D. Perrin
Received June 1984
Revised November 1984

Abstract. A word is called kth power-free if it does not contain any non-empty factor u^k. A morphism is kth power-free if it preserves kth power-free words. A morphism is power-free if it is kth power-free for every $k > 1$.

We show that it is decidable whether a morphism is power-free; more precisely, we prove that a morphism h is power-free iff: h is a square-free morphism and, for each letter a, the image $h(a^2)$ is cube-free.

Introduction

The notion of kth power-free words (i.e., words containing no factor of the form u^k with $u \neq 1$) has been the subject of several works since Thue's paper [8].

An account of basic results may be found in [6, 7]. Berstel [3] gives a survey of some recent results about square-free words and related topics. Properties of kth power-free morphisms and of power-free morphisms are investigated in [1], where the more general concept of an avoidable pattern is introduced.

Usually infinite kth power-free words are constructed by iterating special morphisms. This naturally leads to the notion of kth power-free morphisms (ie., morphisms which preserve the kth power-free property).

For $k = 2$, the decidability of kth power-free property for morphisms was proved in [2]. The characterization of square-free morphisms has been sharpened in [4, 5] and is now optimal.

On the other hand, Bean et al. [1] study, among others things, what we will call here *power-free morphisms*. These are morphisms which preserve kth power-free words for every $k > 1$.

Here we give an effective and simple characterization of power-free morphisms (Theorem 2.1). This result is obtained as a consequence of another result (Theorem 2.2) which shows the relationship between square-free morphisms and kth power-free morphisms.

Section 1 presents some technical lemmas about morphisms which preserve the square-free property of words of length three. Section 2 gives the announced theorems about power-free morphisms and square-free morphisms.

1. Preliminaries

Given a finite alphabet A, we denote by A^* (respectively A^+) the free monoid (respectively semigroup) generated by A. The empty word is denoted by 1, thus $A^+ = A^* - 1$.

A *k-th power* is a nonempty word of the form u^k.

A word is kth power-free if none of its factors is a kth power. If $k=2$ ($k=3$) we say square (cube) instead of kth power. A morphism is a *k-th power-free morphism* provided $h(w)$ is a kth power-free word whenever w is kth power-free.

A morphism is *power-free* if it is a kth power-free morphism for every $k \geq 2$.

A word w is said *primitive* if it is not a proper power of another word (i.e., $w \in u^+$ implies that $w = u$), otherwise w is said *unprimitive.*

The following statement concerning unprimitive words is well known (see, for example, [6]).

Proposition 1.1. *A non-empty word w is unprimitive iff $w = uv = vu$ for some nonempty words u, v.*

We now turn to the study of special morphisms.

Proposition 1.2. *Let h be a morphism from A^* into B^* such that $h(A) \neq \{1\}$. Assume further that $h(w)$ is square-free whenever w is a square-free word of length ≤ 3. Then $h(A)$ is a biprefix code.*

Proof. Let a be a letter of A. If $h(a)=1$, let $b \in A$ with $h(b) \neq 1$; then $h(bab)$ contains a square. Thus $h(a) \neq 1$.

Suppose now that $h(a) \neq 1$ and $h(a)$ is a prefix (respectively suffix) of $h(b)$; then, $h(ab)$ (respectively $h(ba)$) contains a square; a contradiction. □

Lemma 1.3. *Let h be a morphism from A^* into B^* such that $h(w)$ is square-free whenever w is a square-free word of length ≤ 3. Let e_1, $e_2 \in A$ be two letters, and let $v \in A^*$ be a word. Let $h(e_i) = E_i'E_i''$ $(i = 1, 2)$ be factorizations of $h(e_i)$ such that $E_1''E_2' \neq 1$. Assume finally that $E_1''h(v)E_2'$ is a prefix or a suffix of $h(e_0)$ for a letter $e_0 \in A$. Then $v = 1$.*

Proof. By symmetry we consider only the case $E_1''h(v)E_2' = E_0'$ with $h(e_0) = E_0'E_0''$ (see Fig. 1).

Arguing by contradiction, suppose that $v \neq 1$ and set $v = ev'$ with $e \in A$.

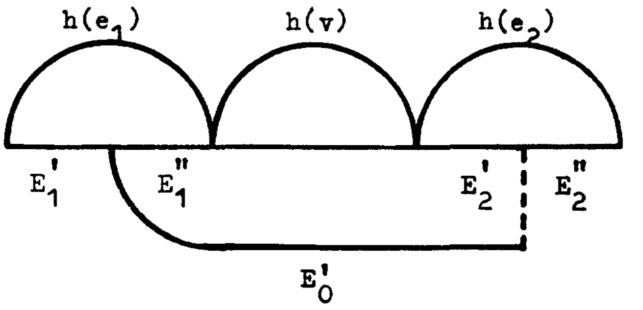

Fig. 1. $E''_1h(v)E'_2$ is a prefix of $h(e_0)$.

Note first that $E''_1 \neq 1$; indeed, on the contrary one would have $E'_2 = 1$, since $h(A)$ is a prefix code, contradicting the hypothesis. Then $h(e_1e_0) = E'_1(E''_1)^2h(v)E'_2E''_0$ contains a square, hence $e_1 = e_0$. It follows that $h(e_0ee_0) = h(e_1ee_0) = E'_1(E''_1h(e))^2h(v')E'_2E''_0$ also contains a square, and consequently $e_0 = e$.

Thus, $h(e_0) = E'_0E''_0 = E''_1h(e_0v')E'_2E''_0$ and $E''_1E'_2 \neq 1$ implies that $h(e_0)$ is a proper factor of itself, which yields the contradiction. □

Proposition 1.4. *Let h be a morphism from A^* into B^* such that $h(w)$ is square-free whenever w is a square-free word of length ≤ 3. Let w, v be two words of A^* such that $h(w) = xh(v)y$ with x, $y \notin h(A^*)$.*

Then there exist a letter $a \in A$ and two words w_1, w_2 of A^* such that $w = w_1aw_2$ and $h(a) = x_1h(v)x_2$, $x = h(w_1)x_1$, $y = x_2h(w_2)$.

Proof. Assume the conclusion is false. There is a letter e of v such that $h(e) = E''_1h(u)E'_2$ where e_1ue_2 is a factor of w with e_1, $e_2 \in A$, $u \in A^*$, and $h(e_i) = E'_iE''_i$ $(i = 1, 2)$.

Note that E''_1 and E'_2 are nonempty words since $h(A)$ is a biprefix code and x, $y \notin h(A^*)$. By Lemma 1.3 we obtain that $u = 1$.

On the other hand, $h(e_1e)$ contains E''^2_1 and $h(ee_2)$ contains E'^2_2. Thus we have $e_1 = e = e_2$.

From $h(e) = E'_1E''_1 = E''_1E'_2 = E'_2E''_2$ we derive that $h(e) = E'_2E''_1 = E''_1E'_2$ since $|E'_1| = |E'_2|$. This means that $h(e)$ is unprimitive and thus $h(e)$ contains a square. This yields the contradiction and completes the proof. □

At last we state the following lemma.

Lemma 1.5. *Let h be a morphism for A^* into B^* such that $h(w)$ is square-free whenever w is a square-free word of length ≤ 3. Let e_i $(i = 1, 2, 3, 4)$ be letters of A and v, $\bar{v}$ be two words of A^*, with $v\bar{v} \neq 1$.*

Assume that $E''_1h(v)E'_2 = E''_3h(\bar{v})E'_4$ with $h(e_i) = E'_iE''_i$ $(i = 1, 2, 3, 4)$ for some factorisations such that E'_2, E'_4 are nonempty words.

Then $E'_2 = E'_4$.

Proof. Assume the contrary. By symmetry it suffices to consider the case $|E'_2| < |E'_4|$.

According to Proposition 1.4 we have that $h(v)$ is a factor of $h(e_4)$ since $|E'_2|<|E'_4|$. Consequently, $E'_4=\bar{E}''_1 h(v)E'_2$ where $\bar{E}''_1$ is some suffix of E''_1.

A first application of Lemma 1.3 gives us $v=1$. Hence $E''_1E'_2=E''_3h(\bar{v})E'_4$ and more precisely $E''_1=E''_3h(\bar{v})\bar{E}'_4$ where $\bar{E}'_4=\bar{E}''_1$ is a prefix of E'_4.

A second application of Lemma 1.3 gives us $\bar{v}=1$. Thus, $v\bar{v}=1$ and this contradicts the assumptions of the lemma. □

2. Power-free morphisms

This section is devoted to an effective characterization of power-free morphisms. That is, we shall prove the following theorem.

Theorem 2.1. *A morphism h is power-free iff h is a square-free morphism and $h(a^2)$ is cube-free for each letter a.*

For a morphism h let us define the *deviation* $e(h)$ of h by

$$e(h)=\max\{|u| \mid h(u) \text{ is a proper factor of } h(e) \text{ for a letter } e\}.$$

This is closely related with the notion of the so-called deviation introduced in [2].

Theorem 2.1 is an immediate consequence of the next theorem. Effectiveness of characterization (we only consider finite alphabets) is shown by condition (iii) which has been proved independently in [4] and [5].

Theorem 2.2. *Let h be a morphism from A^* into B^* such that $h(w)$ is square-free whenever w is a square-free word of length $\leqslant 3$. Then the following conditions hold:*

(i) *h is k-th power-free for all $k>3$.*

(ii) *if $h(a^2)$ is cube-free for each letter $a\in A$, then h is cube-free.*

(iii) *if $h(w)$ is square-free whenever w is a square-free word of length $\leqslant e(h)+2$, then h is square-free.*

Proof of Theorem 2.2. Let w be a word such that $h(w)$ is not kth power-free with $k>1$. Then $|w|\geqslant 2$ since each letter is square-free by hypothesis. We set $w=e_1\ldots e_n$ $(e_i\in A)$. By shortening w if necessary we can assume that $h(w)=E'_1u^kE''_n$ where E''_1, u, E'_n are nonempty words and $h(e_1)=E'_1E''_1$, $h(e_n)=E'_nE''_n$ for some factorizations.

Let us define the *growing sequence* (i_s), $0\leqslant s\leqslant k$, by: $h(e_1\ldots e_{i_s})=E'_1u^sE''_{i_s}$ where $h(e_{i_s})=E'_{i_s}E''_{i_s}$ and $E'_{i_s}\neq 1$ if $s\neq 0$.

Since $h(w)=E'_1u^kE''_n$ and $E'_n\neq 1$ we have $i_0=1$ and $i_k=n$. Now we prove the followings claims.

Claim 2.2.1. *If $1=i_{k-1}$ or $i_1=n$, then $k=2$ and $|w|\leqslant e(h)+2$.*

Proof. By symmetry we suppose $i_{k-1}=1$. By definition of the sequence (i_s), $h(e_1)$ contains a $(k-1)$st power. Hence, $k=2$ and $i_1=1$. Then, $u=E''_{i_1}h(e_2\ldots e_{n-1})E'_n$ and $h(e_1)=E'_1uE''_{i_1}$. Thus, $h(e_2\ldots e_{n-1})$ is a factor of $h(e_1)$ which implies $|e_2\ldots e_{n-1}|\leq e(h)$. Therefore, $|w|\leq e(h)+2$. □

Claim 2.2.2. *If* $i_1=i_{k-1}=2$ *and* $n=3$, *then* w *is not k-th power-free.*

Proof. Since $i_1=i_{k-1}=2$, all factors u from the second up to the $(k-1)$st 'lie' in $h(e_2)$; thus, u^{k-2} is a factor of $h(e_2)$, and consequently $k\leq 3$. Let $w=e_1e_2e_3$. We have $u^{k-1}=E''_1E'_2=E''_2E'_3$ with $h(e_i)=E'_1E''_i$ $(i=1, 2, 3)$.

If $k=2$, then u^2 is a factor of w; hence w is not square-free since $|w|=3$. If $k=3$, then u^2 is a factor of $h(e_1e_2)$ and of $h(e_2e_3)$. Hence, $e_1=e_2$, $e_2=e_3$ and $w=e_1^3$ is not cube-free. □

Claim 2.2.3. *If* $i_1<i_{k-1}$ *and* $n=2$, *then* $w=e^2$ *and* $k=3$.

Proof. Let $w=e_1e_2$. We have $i_1=1$, $i_{k-1}=2$, and $k\geq 3$ since $i_1<i_{k-1}$. If $k>3$, then $i_1<i_{k-2}$ since $h(e_1)$ is square-free. But then $h(e_2)$ is not square-free since $2=i_k=i_{k-1}=i_{k-2}$. Consequently, $k=3$: u^3 is a factor of $h(e_1e_2)$, thus $h(e_1e_2)$ is not square-free, and hence $e_1=e_2$. □

Proof of Theorem 2.2 (*continued*). If $1=i_{k-1}$ or $i_1=n$ we apply Claim 2.2.1. If $i_{k-1}=2$ and $i_1=n-1$ we apply Claim 2.2.2 or Claim 2.2.3 according to whether $i_1=i_{k-1}$ or $i_1<i_{k-1}$. Thus, we can assume that $1<i_{k-1}$, $i_1<n$, and $(2<i_{k-1}$ or $i_1<n-1)$. We have

$$u^{k-1}=E''_1h(e_2\ldots e_{i_{k-1}-1})E'_{i_{k-1}}=E''_{i_1}h(e_{i_1+1}\ldots e_{n-1})E'_n.$$

By construction $E'_{i_{k-1}}$, E'_n are nonempty words and by applying Lemma 1.5 on factorizations of u^{k-1} we obtain $E'_{i_{k-1}}=E'_n$. Since $h(A)$ is a biprefix code, this implies (see Fig. 2) for all j, t with $0<j<i_1$, $0<t<k$, the equalities

$$e_{i_t+j}=e_j \quad\text{and}\quad e_{i_t}=e_n \qquad \text{if } E''_n=1.$$

$$e_{i_t+j}=e_{j+1},\quad E''_1=E''_{i_t}\quad\text{and}\quad E'_{i_t}=E'_n \qquad \text{if } E''_n\neq 1.$$

The asymmetry of these formulas is due to the fact that $E''_1\neq 1$.

We deduce from them that $w=(e_1\ldots e_{i_1})^k$ if $E''_n=1$ and that $h(e_1e_{i_1}e_n)=E'_1(E''_1E'_{i_1})^2E'_n$ if $E''_n\neq 1$.

In the second case, $e_n=e_{i_1}$ (or $e_{i_1}=e_1$) since $h(e_1e_{i_1}e_n)$ contains a square, and hence $w=e_1\ (e_2\ldots e_{i_1})^k$ (or $w=(e_1\ldots e_{i_1-1})^ke_n$). Thus w is not kth power-free and this completes the proof. □

The condition that $h(a^2)$ is cube-free for each letter a is necessary, as is shown by the following example due to Bean et al. [1].

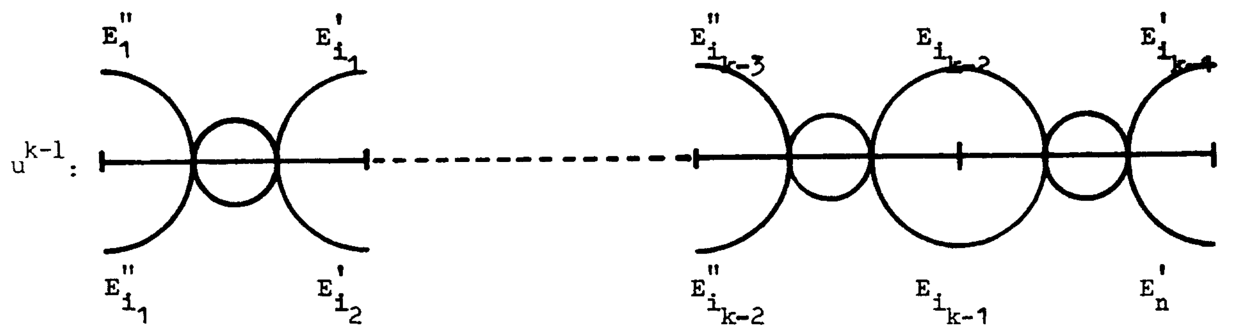

Fig. 2. $u^{k-1} = E''_1 h(e_2 \dots e_{i_{k-1}-1}) E'_{i_{k-1}} = E''_{i_1} h(e_{i_1+1} \dots e_{n-1}) E'_n$.

Example. Let h be an endomorphism on $\{a, b, c, d\}^*$ induced by

$$a \vdash abacbab, \quad c \vdash cdacabcbd,$$

$$b \vdash cdabcabd, \quad d \vdash cdacbcacbd.$$

h is square-free according to Theorem 2.2 ($e(h) = 0$) but $h(a^2) = abac(ba)^3cbab$ and so h is not power-free.

Acknowledgment

I would like to thank Professors J. Berstel and D. Perrin for their encouragements.

References

[1] D. Bean, A. Ehrenfeucht and G. McNulty, Avoidable patterns in strings of symbols, *Pacific J. Math.* **85** (2) (1979) 261–294.
[2] J. Berstel, Sur les mots sans carré définis par un morphisme, In: *6th ICALP Symp.*, Lecture Notes in Computer Science **71** (Springer, Berlin, 1979) 16–25.
[3] J. Berstel, Some recent results on square-free words (STACS' 84), Tech. Rept. LITP No. 84-6, 1984.
[4] M. Crochemore, Sharp characterizations of square-free morphisms, *Theoret. Comput. Sci.* **18** (1982) 221–226.
[5] A. Ehrenfeucht and G. Rozenberg, Repetitions in homomorphisms and languages, in: *9th ICALP Symp.*, Lecture Notes in Computer Science **140** (Springer, Berlin, 1982) 192–196.
[6] M. Lothaire, Combinatoric on words, in: G.-C. Rota, ed., *Encyclopedia of Mathematics and its Applications Vol. 17* (Addison-Wesley, Reading, MA, 1983).
[7] A. Salomaa, *Jewels of Formal Language Theory* (Pitman, London, 1981).
[8] A. Thue, Über unendliche Zeichenreihen, *Norske Vid. Selsk. Skr. Mat. Nat. Kl* (*Kristiana*) 7 (1906) 1–22.