INFORMATIQUE THÉORIQUE ET APPLICATIONS

TERO HARJU
MATTI LINNA

On the periodicity of morphisms on free monoids

Informatique théorique et applications, tome 20, nº 1 (1986), p. 47-54
<http://www.numdam.org/item?id=ITA_1986__20_1_47_0>

Informatique théorique et Applications/Theoretical Informatics and Applications
(vol. 20, n° 1, 1986, p. 47 à 54)

ON THE PERIODICITY OF MORPHISMS ON FREE MONOIDS (*)

by Tero HARJU [1] and Matti LINNA [1]

Communicated by J. BERSTEL

Abstract. — *For a given endomorphism h on a finitely generated free monoid there are only finitely many primitive words w for which* $h(w)=w^n$ *for some* $n \geq 2$. *Also, one can effectively find all such w. Using this it is shown that it is decidable whether or not a morphism h defines an ultimately periodic infinite word when iterated on a given word. This latter result thus solves the DOL periodicity problem.*

Résumé. — *Soit A un alphabet fini, A* le monoïde libre engendré par A. Pour tout endomorphisme h de A* il existe seulement un nombre fini de mots primitifs w tels que* $h(w)=w^n$ *avec un entier* $n \geq 2$. *Aussi peut-on effectivement trouver tous les w de cette espèce. Avec cela on prouve qu'on peut résoudre effectivement le problème suivant : Est-ce que le mot infini obtenu par itération de h est ultimement périodique. Cela résout le problème de périodicité des DOL systèmes.*

1. INTRODUCTION

We shall first consider equations $h(x)=x^n$, $n=2, 3, \ldots$, where h is an endomorphism on a finitely generated free monoid A^*. It is shown that all the solutions of these are obtained as powers of finitely many primitive words. The finiteness of the primitive solutions was first shown in [6]. Also, we prove that the primitive solutions can be effectively determined. This, in particular, shows that it is decidable whether or not a nontrivial solution exists.

In chapter 4 we turn to infinite words obtained by iterating an endomorphism $h: A^* \to A^*$. We show that the DOL periodicity problem is decidable, that is, we can decide whether there exist words v and w for a given word u such that $h^{\omega}(u)=vw^{\omega}$, where $h^{\omega}(u)$ is the limit of the sequence u, $h(u)$,

(*) Received April 1984, revised March 1985.
(1) Department of Mathematics, University of Turku, SF-20500 Turku, Finland.

ISSN en cours 86/01 478/$ 2.80/© Gauthier-Villars.

$h^2(u), \ldots$ if it exists and $w^\omega = ww \ldots$. Some special cases of this problem were solved in [4] and [6]. In [4] a partial solution to the problem was used to solve the so called adherence equivalence problem for DOL systems.

The ultimate periodicity problem comes into use also in solving the ω-regularity problem for limits of DOL languages. The ordinary regularity problem for DOL languages was shown to be decidable in [7]. The corresponding problem for infinite words is just another formulation for the DOL periodicity problem and is thus solved here.

2. PRELIMINARIES

Let A be a finite alphabet and A^* the free monoid generated by A. We denote by 1 the identity (the empty word) in A^* and by A^+ the free semigroup $A^* \setminus \{1\}$. For a word $w \in A^*$, $|w|$ denotes the length of w, while $|A|$ is the cardinality of A.

A word $w \in A^*$ is *primitive* if it is not a power of another word. Every word is a power of a primitive word, denoted by $\sqrt{w}$. Given two words w and v we say that w is a *prefix* of v, $w \leqq v$, in case $v = ww_1$ for some $w_1 \in A^*$. Also, w and v are *conjugates*, $w \approx v$, if one finds words u_1 and u_2 such that $w = u_1 u_2$ and $v = u_2 u_1$. We may write in this case $v = u_1^{-1} w u_1$ with the obvious meaning.

Given subsets $B_1, \ldots, B_r$ of A^* the wording u *decomposes as* $u = u_1 \ldots u_r$ *in* $B_1 \ldots B_r$ means that $u_i \in B_i$.

In what follows we are interested in iterating a morphism $h: A^* \to A^*$ starting with a given word u. This iteration gives us a sequence

$$u, h(u), h^2(u), \ldots \tag{2.1}$$

of words. The pair (h, u) is called a *DOL system* in the literature and its language is the set $L(h, u) = \{h^i(u) \mid i \geqq 0\}$, which may be finite, of course.

The limit, *lim* $L(h, u)$, of the set $L(h, u)$ consists of all *infinite* words $\alpha = a_1 a_2 \ldots, a_i \in A$, such that for all n, α *possesses a prefix longer than* n belonging to (2.1).

The question whether or not $\lim L(h, u) \neq \varphi$ was shown to be decidable in [5]. Moreover, if this limit is nonempty then one can effectively find integers p and q such that $h^p(u)$ is a proper prefix of $h^{p+q}(u)$. Also, in this case:

$$\lim L(h, u) = \bigcup_{i=0}^{q-1} \lim L(h^q, h^{p+i}(u)),$$

where moreover $|\lim L(h^q, h^{p+i}(u))| = 1$ for each $i = 0, \ldots, q-1$. We shall thus separate the case (2.1) in an effective way into a finite number of special cases where the limit of the sequence exists uniquely.

As discussed above, we can restrict ourselves to DOL systems (h, u), where (if the limit exists) $h(u) = ux$ for some $x \in A^*$. This kind of a system defines the infinite word:

$$h^\omega(u) = uxh(x)h^2(x)\ldots$$

In chapter 4 we shall show that it is decidable whether or not such a prefix preserving h defines an *ultimately periodic* infinite word, that is, whether or not:

$$h^\omega(u) = vw^\omega$$

for some words v and w. Here w^ω denotes the infinite word $ww\ldots$

Given a morphism $h: A^* \to A^*$ we call a letter $b \in A$ *finite* if $L(h, b)$ is a finite set. Otherwise b is an *infinite letter*. This offers a partition $A = A_F \cup A_I$ for A. Finally, if $b \in A$ is "mortal", that is, $h^i(b) = 1$ for some $i \geqq 0$ then necessarily $h^{|A|}(b) = 1$.

3. THE EQUATIONS $h(x) = x^n$

In this chapter we shall seek for the solutions of the equations:

$$h(x) = x^n, \qquad n = 2, 3, \ldots, \tag{3.1}$$

where $h: A^* \to A^*$ is a given morphism. We shall show that all the solutions can be effectively found.

Given a solution, $h(w) = w^n$ for some $n \geqq 2$, we note that $(\sqrt{w})^p$ is also a solution for all $p \geqq 0$. Thus we need to search for the primitive solutions only. With this in mind we define:

$$P_h = \{ w \in A^+ \mid w \text{ primitive and } h(w) = w^n \text{ for some } n \geqq 2 \}.$$

Let $A = A_F \cup A_I$, where A_F is the set of finite letters and A_I the set of infinite letters with respect to h.

THEOREM 1 : *For a given $h: A^* \to A^*$ there are only finitely many primitive words w for which $h(w) = w^n$ for some $n \geqq 2$. In fact there is a partition*

$A_1, \ldots, A_r$ of A_I such that:

$$P_h \subseteq \bigcup_{i=1}^{r} (A_F \cup A_i)^*$$

and the words in $P_h \cap (A_F \cup A_i)^*$ are conjugates ($i=1, \ldots, r$).

Proof: We note first that every solution contains an infinite letter. Let w and v be two primitive solutions: $h(w)=w^n$ and $h(v)=v^m$ (n, $m \geqq 2$) and assume these share an infinite letter, say b. There is then an integer k such that $|h^k(b)| \geqq 2.\max\{|w|, |v|\}$. Now since $h^\omega(w)=w^\omega$, $h^\omega(v)=v^\omega$ and b occurs in both w and v, there are words $w_1 \approx w$ and $v_1 \approx v$ so that:

$$h^k(b)=w_1^i w_2 = v_1^j v_2,$$

where i, $j \geqq 2$ and $w_2 < w_1$, $v_2 < v_1$. This means that $w_1 = v_1$ since they are primitive. In all $w \approx v$ and the claim follows.

We now proceed to show that the primitive solutions can be found in an effective way.

THEOREM 2: *The set P_h can be constructed effectively for given $h: A^* \to A^*$.*

Proof: For each infinite letter b we shall decide whether or not b is the leftmost infinite letter of a solution and if so then we determine this solution. By the previous theorem the present claim then follows.

Let $b \in A_I$ be given. Then $h(b)$ has a decomposition

$$h(b)=xbu$$

in $A_F^* A_I A^*$, where $h^{|A|}(x)=1$. Otherwise b cannot be as required. Now h generates an infinite word:

$$h^\omega(b)=u_0 bu_1 bu_2 \ldots,$$

where $u_i \in (A \setminus \{b\})^*$ and $h(u_0 b)=u_0 bu$. If here the set $U=\{u_i \mid i \geqq 0\}$ is infinite then b cannot occur in any solution. We remark first that if an infinite letter c occurs in $h^\omega(b)$ then it occurs in $h^i(b)$ for some $i \leqq |A|$. If for such a letter c the letter b occurs in no $h^i(c)$ then U is infinite. This property is clearly a decidable one and so we assume that it does not take place.

CLAIM: *U is infinite iff there is an infinite letter c in $h^\omega(b)$ and an integer $s \leqq |A|$ such that $h^s(c)=v_1 cv_2$ and for $i=1$ or $2: v_i \in A_F^+$ and $h^{|A|}(v_i) \neq 1$.*

Proof of the claim: Assume there are arbitrarily long words in U. Since each infinite letter from $h^\omega(b)$ produces an occurrence of b in at most $|A|$

steps there are arbitrarily long words from A_F^* as subwords in U. This is possible only as stated in the claim. The converse is trivial.

By this claim it is decidable whether or not U is finite and if so then U can be constructed easily. We suppose then that U is finite and given.

Assume for a while that there is a solution w, $h(w)=w^n$, such that w decomposes as $w=w_0 u_0 b w_1$ in $A_F^* A_F^* A_I A^*$. Then:

$$w^n = h(w) = h(w_0) u_0 buh(w_1)$$

and so $w_0 = h(w_0)$. We consider the conjugate $v = w_0^{-1} w w_0$. We have:

$$h(v) = u_0 buh(w_1) w_0 = w_0^{-1} w^n w_0$$

and thus:

$$(3.2) \qquad h(v) = v^n \text{ and } v \in A_F^* A_I A^*.$$

In particular we have $h^\omega(b) = v^\omega$ and thus there exists an integer k with $h^k(b) \in v A^*$. Since v is primitive:

$$(3.3) \qquad h^k(bu_i) \in v^* \quad \text{for all } u_i \in U.$$

After this we omit the assumption for the existence of such a solution.

Given two words v_1 and v_2 it is decidable whether or not $h^i(v_1) = h^i(v_2)$ for all $i \geqq m$ for some m, and if so then m can be found effectively, [1 or 2]. We apply this effectiveness result to the starting words $bu_i bu_j$ and $bu_j bu_i$ for all u_i, $u_j \in U$. If some of these pairs are not ultimately equivalent then (3.2) cannot hold for any v because of (3.3). So suppose m is now an effective integer such that:

$$h^m(bu_i bu_j) = h^m(bu_j bu_i) \quad \text{for all } u_i, u_j \in U.$$

Because of this commutation and transitivity:

$$h^1(bu_i) \in z^* \quad \text{for all } u_i \in U, \quad 1 \geqq m,$$

where z is a primitive word effectively obtainable. However, if (3.2) holds then necessarily $v = z$. This proves the theorem.

The following decidability result is an immediate consequence.

COROLLARY 3: *It is decidable whether or not the equations* $h(x) = x^n$, $n = 2, 3, \ldots$ *possess a nontrivial solution.*

We close this chapter by giving a somewhat stronger theorem resulting from the above.

THEOREM 4: *For a given h it is decidable whether or not there exists a nontrivial word w such that $h^m(w)=w^n$ for some $m \geqq 1$ and $n \geqq 2$.*

Proof: Suppose $h^m(w)=w^n$ for some $m \geqq |A|$ and $n \geqq 2$. It suffices to show that there are integers $p<|A|$ and $q \geqq 2$ such that $h^p(w)=w^q$. As in the proof of theorem 2 we may assume that w decomposes as $w=u_0 bw_1$ in $A_F^* A_I A^*$ and $h^m(b)=u_0 bu$ for some $u \in A$. Now there exists an integer $p<|A|$ such that $h^p(b)=u_0 bu_1$ for some $u_1 \in A^*$. Since $(h^p)^\omega (u_0 b)=(h^m)^\omega (u_0 b)=w^\omega$ the claim follows.

4. ULTIMATE PERIODICITY

We shall use the results from the previous chapter in proving the decidability of the ultimate periodicity problem. We remind that the morphism h is ultimately periodic on a word u if:

$$h^\omega(u)=vw^\omega, \tag{4.1}$$

for some words v and w. Here we may suppose that $h(u)=ux$ as pointed out in preliminaries.

THEOREM 5: *The ultimate periodicity problem is decidable for DOL systems.*

Proof: Let us be given a morphism $h: A^* \to A^*$ and a word $u \in A^*$ such that $h(u)=ux$ for some $x \in A^+$. Denote by A_1 the subset of A which consists of the infinite letters occuring infinitely many times in $h^\omega(u)$. Clearly A_1 is an effective set.

In case $A_1=\varphi$ there appears only one infinite letter b which is isolated, that is, no letter produces b. This case is thus easy, since the period comes out from $h(b)=u_0 bv$.

Assume now that $A_1 \neq \varphi$ and let b be the first letter of $h^\omega(u)$ from A_1. Then for some $i \leqq |A|$ *and* $y \in A^*$:

$$h^i(yb)=yby_1 \qquad \text{and} \qquad h^{|A|}(y)=1.$$

For otherwise $h^\omega(u)$ is not ultimately periodic. We may assume that $i=1$ since otherwise we consider the pair (h^i, u) instead of (h, u). Thus:

$$h(yb)=yby_1 \qquad \text{and} \qquad h^{|A|}(y)=1.$$

Let us write now:

$$h^\omega(u)=u_1\, ybu_2\, u_3 \ldots,$$

where:

$$h(u_1)=u_1\,ybu_2 \tag{4.2}$$

and u_3 is any (but fixed) factor such that:

$$u_3 \in A^* A_I A^*. \tag{4.3}$$

By theorem 2 we may test whether there exists a primitive w such that $yb \leqq w$ and $h(w)=w^n$ for some $n \geqq 2$. If no such w can be found then $h^\omega(u)$ is not ultimately periodic by above. Assume then that we have found such a word w.

CLAIM: *$h^\omega(u)$ is ultimately periodic iff $h^\omega(ybu_2\,u_3)=w^\omega$ iff $(h, ybu_2\,u_3)$ defines an infinite word.*

Proof of the claim: Assume $h^\omega(ybu_2\,u_3)$ is defined. Then:

$$h^\omega(ybu_2\,u_3)=h^\omega(yb)=w^\omega \qquad (\text{since } b \in A_I \text{ and } yb \leqq w).$$

We have also for all $i \geqq 1$:

$$h^i(u_1\,ybu_2\,u_3)=u_1 \,.\, ybu_2 \,.\,.\,.\,.\,.\, h^i(ybu_2) \,.\, h^i(u_3)$$

and since $h^i(u_1\,ybu_2\,u_3)$ is a prefix of $h^{i+1}(u_1\,ybu_2\,u_3)$:

$$h^i(u_3) \leqq h^{i+1}(ybu_2\,u_3). \tag{4.4}$$

Suppose i is here already so large that $|h^i(u_3)|>|w|$ and $h^j(ybu_2\,u_3) \leqq w^\omega$ for all $j \geqq i$. Then by (4.4), $w \leqq h^i(u_3)$ and hence $h^i(ybu_2) \in w^*$. Now $h^\omega(u)=h^\omega(u_1\,ybu_2\,u_3)$ implies that $h^\omega(u)$ is ultimately periodic. The converse of the claim is trivial.

We note that the last statement is decidable in the claim and so is the first one. This completes the proof of the theorem.

5. DISCUSSION

Theorems 1 and 2 or their proofs do not give the primitive solutions w explicitely. In the binary case, $|A|=2$, one can, however, obtain a very effective characterization to the set P_h as well as to the morphism h, [3]:

THEOREM 6: *Let w be a primitive word in $\{a, b\}^*$ and let h be an endomorphism on $\{a, b\}^*$. Then $h(w)=w^n$ for some $n \geqq 2$ iff at least one of the letters, say a, is infinite and:*

(1) *$w=\sqrt{h}(a)$ and either:*

(i) *$h(b)=1$ and, $|w|_a \geqq 2$ or $h(a)$ is not primitive;*

or

(ii) $w=\sqrt{h}(b)$;

or

(iii) $h(a)\in a^2 a^*$;

or

(2) $w=b^{s_1} ab^{s_2}$ *and* $h(b)=b$, $h(a)\in(ab^{s_1+s_2})^+ a$;

or

(3) $w=ab$ *and* $h(a)\in(ab)^+a$, $h(b)\in b(ab)^+$.

The finiteness result in theorem 1 is characteristic to free semigroups. Namely, this property fails already in *one-relator semigroups*. As an example we mention the free commutative semigroup $\langle$ a, b; $ab=ba\rangle$, where the morphism h defined as $h(a)=a^2$, $h(b)=b^2$ possesses infinitely many 'primitive' solutions of the form $a^i b$.

Also it is worthwhile noting that this failure concerns *free groups* as well. To see this consider a word w in a finitely generated free group F such that $h(w)=w^n$ for an endormorphism h on F. Now, let G be a disjoint finitely generated free group and define h to be the identity on G. We have $h(uwu^{-1})=uw^n u^{-1}=(uwu^{-1})^n$ for all u in G and so h has infinitely many solutions of the form uwu^{-1} in the free product $G^* F$.

REFERENCES

1. K. Culik II, *The Ultimate Equivalence Problem for DOL Systems*, Acta Informatica, Vol. 10, 1978, pp. 79-84.
2. A. Ehrenfeucht and G. Rozenberg, *On Simplifications of PDOL Systems*, Proceedings of a Conference on Theoretical Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 1977.
3. T. Harju and M. Linna, *The equations $h(w)=w^n$ in binary alphabets*, Theoret. Comput. Sci. 33, 1984, pp. 327-329.
4. T. Head, *Adherence Equivalence is Decidable for DOL Languages*, Proceedings of a Symposium of Theoretical Aspects of Computer Science, Paris, 1984, Springer Lecture Notes in Computer Science, Vol. 166, 1984, pp. 241-248.
5. M. Linna, *The Decidability of the DOL Prefix Problem*, Int. J. Comput. Math., Vol. 6, 1977, pp. 127-142.
6. M. Linna, *On Periodic ω-Sequences Obtained by Iterating Morphisms*, Ann. Univ. Turkuensis, Ser. A I 186, 1984, pp. 64-71.
7. A. Salomaa, *Comparative Decision Problems between Sequential and Parallel Rewriting*, Proc. Symp. Uniformly Structured Automata and Logic, Tokyo, 1975, pp. 62-66.