

# On the Periodicity of Morphic Words

Vesa Halava<sup>1</sup>, Tero Harju<sup>1</sup>, Tomi Kärki<sup>1\*</sup>, and Michel Rigo<sup>2</sup>

<sup>1</sup> Department of Mathematics, University of Turku,  
FI-20014 Turku, Finland  
{vehalava,harju,topeka}@utu.fi

<sup>2</sup> Institute of Mathematics, University of Liège,  
Grande Traverse 12 (B 37), B-4000 Liège, Belgium  
M.Rigo@ulg.ac.be

**Abstract.** Given a morphism  $h$  prolongable on  $a$  and an integer  $p$ , we present an algorithm that calculates which letters occur infinitely often in congruent positions modulo  $p$  in the infinite word  $h^\omega(a)$ . As a corollary, we show that it is decidable whether a morphic word is ultimately  $p$ -periodic. Moreover, using our algorithm we can find the smallest similarity relation such that the morphic word is ultimately relationally  $p$ -periodic. The problem of deciding whether an automatic sequence is ultimately weakly  $R$ -periodic is also shown to be decidable.

**Key words:** automatic sequence, decidability, morphic word, periodicity, similarity relation

## 1 Introduction

Periodicity is one of the main concepts in combinatorics on words and an important subject also from the application point of view [11]. The most famous periodicity results in word combinatorics are probably the theorem of Fine and Wilf, the Critical Factorization Theorem, and the theorem of Morse and Hedlund. The theorem of Fine and Wilf considers two simultaneously occurring periods in one word [6], the Critical Factorization Theorem relates the period of a word with the local repetitions [3, 5], and the theorem of Morse and Hedlund characterizes ultimately periodic words in terms of subword complexity [9].

It is an interesting and important question how to recognize periodicity and ultimate periodicity of infinite words. This depends heavily on the way the infinite word is generated. For example, in 1986, J. Honkala proved that it is decidable whether a given automatic sequence is ultimately periodic [10]. Such a sequence, produced by a finite automaton with output, can also be generated using a uniform morphism and a coding. Recently, a new method for solving the ultimate periodicity problem for automatic sequences was given by J.-P. Alouche, N. Rampersad and J. Shallit [1]. A more general result showing the decidability of the ultimate periodicity question for pure morphic words was given by T. Harju and M. Linna [8] and, independently, by J.-J. Pansiot [12].

---

\* Supported by Finnish Cultural Foundation, Varsinais-Suomi Regional fund.

However, if we replace pure morphic words of the form  $h^\omega(a)$  by general morphic words of the form  $g(h^\omega(a))$  where  $g$  is a coding, then the decidability status of the ultimate periodicity problem remains unknown.

In this paper, we solve the decision problem of the ultimate  $p$ -periodicity of morphic words. For given integer  $p$  and morphism  $h$  prolongable on  $a$ , i.e.,  $h(a) = ax$  and  $h^n(x) \neq \varepsilon$  for every  $n \geq 0$ , we present an algorithm that calculates which letters occur infinitely often in congruent positions modulo  $p$  in the given infinite word  $h^\omega(a)$ . It follows from this that, for any coding  $g$ , it is decidable whether  $g(h^\omega(a))$  is ultimately  $p$ -periodic. We then consider relational periodicity of morphic words. Relational periods were introduced in [7] as a generalization of periods in partial words. Using our algorithm we can find the smallest similarity relation such that the infinite word is ultimately relationally  $p$ -periodic. Finally, we show that given a similarity relation  $R$  it is decidable whether an automatic sequence is ultimately weakly  $R$ -periodic.

## 2 Preliminaries

Let  $\mathcal{A}$  be a finite alphabet and denote the empty word by  $\varepsilon$ . The free monoid  $\mathcal{A}^*$  is the set of finite words over  $\mathcal{A}$  with the operation of concatenation. The length of a finite word  $w$  is denoted by  $|w|$ . The set of letters occurring in a word  $w$  is denoted by  $\text{Alph}(w)$ . An infinite word over  $\mathcal{A}$  is a sequence  $x = (x_n)_{n \geq 0}$  where  $x_n \in \mathcal{A}$  for every  $n \geq 0$ . Denote the set of all infinite words over  $\mathcal{A}$  by  $\mathcal{A}^\omega$ . Let also  $\mathcal{A}^\infty = \mathcal{A}^* \cup \mathcal{A}^\omega$ . A word  $v$  is a *factor* of  $w \in \mathcal{A}^\infty$  if  $w = xvy$  for some word  $x \in \mathcal{A}^*$  and  $y \in \mathcal{A}^\infty$ . The factor  $v$  is called a *prefix* of  $w$  if  $x = \varepsilon$  and it is called a *suffix* of  $w$  if  $y = \varepsilon$ . The set of factors of length  $n$  of a word  $w$  is denoted by  $F_n(w)$ .

A mapping  $h: \mathcal{A}^* \rightarrow \mathcal{A}^*$  is a morphism if  $h(uv) = h(u)h(v)$  for every  $u, v \in \mathcal{A}^*$ . A morphism  $h$  on  $\mathcal{A}^*$  is *prolongable* on a letter  $a$  if  $h(a) = ay$  and  $h^n(y) \neq \varepsilon$  for all integers  $n \geq 0$ . In this case,  $h^n(a)$  is a prefix of  $h^{n+1}(a)$  and the following fixed point of  $h$  exists:

$$h^\omega(a) = \lim_{n \rightarrow \infty} h^n(a) = ayh(y)h^2(y) \dots$$

An infinite word of the above form  $x = h^\omega(a)$  is called a *pure morphic word*. A *morphic word* is an image of a pure morphic words by a coding, i.e., it is of the form  $g(h^\omega(a))$  for some morphism  $h$  prolongable on  $a$  and for some coding, i.e., a letter-to-letter morphism  $g: \mathcal{A}^* \rightarrow \mathcal{B}^*$ .

An infinite word  $x$  is ultimately periodic if it is of the form  $x = uv^\omega = uvvv \dots$ , where  $u$  and  $v$  are finite words. The length  $|u|$  is a *preperiod* and the length  $|v|$  is a *period* of  $x$ . An infinite word  $x$  is *ultimately  $p$ -periodic* if  $|v| = p$ . The smallest period of  $x$  is called *the period* of  $x$ .

A *similarity relation*  $R$  is a relation on finite words induced by a reflexive and symmetric relation on letters. A word  $u = u_1 \dots u_m$  is  *$R$ -similar* to a word  $v = v_1 \dots v_n$  if  $n = m$  and  $u_i R v_i$  for all letters  $u_i, v_i \in \mathcal{A}$ . Note that a similarity relation need not be transitive. For example, if  $a R c$ ,  $b R c$  and  $(a, b) \notin R$ , then

$abba Rcbca$ , but  $abba$  and  $caca$  are not  $R$ -similar since the second letters are not related.

Three kinds of relational periods with respect to a given similarity relation  $R$  were introduced in [7]. An infinite word  $x = (x_n)_{n \geq 0}$  is *weakly  $R$ -periodic* if  $x_i R x_{i+p}$  for every  $i \geq 0$  and for some integer  $p > 0$ , which is called a *weak  $R$ -period* of  $x$ . It is *externally  $R$ -periodic* with an *external  $R$ -period*  $p > 0$  if there exists a word  $v = v_0 \cdots v_{p-1}$  such that  $x_i R v_j$  if  $i \equiv j \pmod{p}$ . The word  $x$  is *strongly  $R$ -periodic* with a *strong  $R$ -period*  $p > 0$  if  $i \equiv j \pmod{p}$  implies  $x_i R x_j$ . It is shown in [7] that all these relational periods are different, and any strong  $R$ -period of  $x$  is also a weak and an external  $R$ -period. An infinite word  $x$  is ultimately weakly (resp. externally, strongly)  $R$ -periodic if  $x = uv$  where  $u$  is a finite word (a preperiod) and the infinite word  $v$  is weakly (resp. externally, strongly)  $R$ -periodic. We say that an infinite word is *ultimately relationally  $p$ -periodic* if  $p \in \mathbb{N}$  is a weak, external or strong  $R$ -period of some suffix of the word.

Let  $k \geq 2$  be an integer. An infinite word  $x = (x_n)_{n \geq 0} \in \mathcal{B}^\omega$  is  *$k$ -automatic* if there exists a finite deterministic automaton with output  $M = (Q, q_0, \Sigma_k, \delta, \mathcal{B}, \tau)$  such that  $\tau(\delta(q_0, \text{rep}_k(n))) = x_n$  for all  $n \geq 0$ . Here  $\text{rep}_k(n)$  denotes the base- $k$  representation of the integer  $n$ ,  $Q$  is the finite set of states,  $q_0$  is the initial state,  $\Sigma_k = \{0, \dots, k-1\}$  is the input alphabet,  $\delta$  is the transition function and  $\tau: Q \rightarrow \mathcal{B}$  is the output function. By the result of Cobham [4], an infinite word is  $k$ -automatic if and only if it is of the form  $g(h^\omega(a))$  for a  $k$ -uniform morphism  $h$  prolongable on  $a$  and a coding  $g$ . A morphism  $h: \mathcal{A}^* \rightarrow \mathcal{A}^*$  is  *$k$ -uniform* if  $|h(b)| = k$  for all letters  $b \in \mathcal{A}$ .

### 3 Algorithm on $k$ -sets

Let  $x = (x_n)_{n \geq 0}$  be an infinite word over  $\mathcal{A}$ . Let  $k \in \{0, 1, \dots, p-1\}$ . We say that a letter  $a \in \mathcal{A}$  belongs to the  *$k$ -set of  $x$  modulo  $p$*  if there exist infinitely many integers  $n$  such that

$$x_n = a \quad \text{and} \quad n \equiv k \pmod{p}.$$

In this section, we show that given integers  $k$  and  $p$ , it is decidable whether a letter  $a$  belongs to the  $k$ -set of a given purely morphic word  $x$  modulo  $p$ . Consequently, we give an algorithm that counts all  $k$ -sets of  $x$  modulo  $p$ . First, let us prove the following lemma concerning iteration of morphisms.

**Lemma 1.** *Let  $h: \mathcal{A}^* \rightarrow \mathcal{A}^*$  be a morphism, and let  $p$  be a positive integer. There exist integers  $r$  and  $q > 0$  such that*

$$|h^r(b)| \equiv |h^{r+q}(b)| \pmod{p} \tag{1}$$

for all  $b \in \mathcal{A}$ .

*Proof.* If  $p = 1$ , then the claim is trivial. Let  $p > 1$ . Let  $M = M(h)$  be the incidence matrix of the morphism  $h$  on the alphabet  $\mathcal{A} = \{a_1, a_2, \dots, a_d\}$ , i.e.,

$$M = (m_{i,j})_{1 \leq i,j \leq d},$$

where  $m_{i,j}$  denotes the number of occurrences of  $a_i$  in  $h(a_j)$ . Now

$$|h^n(a_j)| = v_t M^n v_j^T,$$

for  $v_t = (1, \dots, 1)$  and  $v_j = (\delta_{ij})_{1 \leq i \leq d}$ , where  $\delta_{jj} = 1$  and  $\delta_{ij} = 0$  for  $i \neq j$ . There are only finitely many matrices  $M^n \bmod p$ , where the entries are the residues modulo  $p$ . Hence the infinite sequence  $(u_{j,n})_{n \geq 0}$  of the lengths  $u_{j,n} = |h^n(a_j)| \bmod p$  is ultimately periodic. Denote the preperiod of  $(u_{j,n})_{n \geq 0}$  by  $r_j$  and the period of  $(u_{j,n})_{n \geq 0}$  by  $q_j$ . Set  $r = \max r_j$  and  $q = \text{lcm}(q_1, \dots, q_d)$ . By the periodicity of the sequences  $(u_{j,n})_{n \geq 0}$ , we conclude that  $|h^r(a_j)| \equiv |h^{r+q}(a_j)| \pmod{p}$  for any  $a_j \in \mathcal{A}$ .  $\square$

Using the above lemma, we now prove our main result.

**Theorem 1.** *Let  $x = (x_n)_{n \geq 0} = h^\omega(a) \in \mathcal{A}^\omega$  for a morphism  $h$  prolongable on  $a$ , and let  $b$  be a letter in  $\mathcal{A}$ . Given positive integers  $k$  and  $p$ , it is decidable whether there exist infinitely many  $n$  such that  $x_n = b$  and  $n \equiv k \pmod{p}$ .*

*Proof.* Let  $r$  and  $q$  be the integers satisfying (1) for the morphism  $h$  and the given integer  $p$ . Consider the directed graph  $G_h = (V, E)$  where the set of vertices  $V$  is  $\{(a, i) \mid a \in \mathcal{A}, 0 \leq i < p\}$  and there is an edge from  $(c, i)$  to  $(d, j)$  if there exist a letter  $b \in \mathcal{A}$  and integers  $m$  and  $m'$  such that

$$h^r(b) = y_1 \cdots y_n \text{ and } y_m = c \text{ with } 1 \leq m \leq n, \quad (2)$$

$$h^q(c) = z_1 \cdots z_{n'} \text{ and } z_{m'} = d \text{ with } 1 \leq m' \leq n', \quad (3)$$

$$j - i \equiv |h^q(y_1 \cdots y_{m-1})| + m' - m \pmod{p}. \quad (4)$$

By Lemma 1, this means that if  $c$  is at position  $i \bmod p$  in  $x$  and it is the  $m$ th letter of the image  $h^r(b)$  for some  $b$  in  $x$ , then there is  $d$  at position  $j \bmod p$  of  $x$  and it is the  $m'$ th letter of the image  $h^q(c)$ ; see Figure 1. Namely, consider the position of  $d$  modulo  $p$ . If  $x_l = b$ , then

$$i \equiv |h^r(x_0 \cdots x_{l-1})| + m - 1 \pmod{p}, \quad (5)$$

$$j \equiv |h^{r+q}(x_0 \cdots x_{l-1})| + |h^q(y_1 \cdots y_{m-1})| + m' - 1 \pmod{p}. \quad (6)$$

By (1), we have  $|h^{r+q}(x_0 \cdots x_{l-1})| \equiv |h^r(x_0 \cdots x_{l-1})| \pmod{p}$ , which together with (5) and (6) implies (4).

Recall that  $h$  is prolongable on  $a$ . We say that a vertex  $(c, i) \in V$  is an *initial vertex* if there exists a letter  $b = x_l$  such that  $0 \leq l < |h^r(a)|$ ,  $c$  is the  $m$ th letter of  $h^r(b)$  and  $i$  satisfies (5). The set of the initial vertices is denoted by  $V_I$ . Moreover, if there exist infinitely many paths in  $G_h$  starting from some vertex of  $V_I$  and ending in  $v \in V$ , we say that  $v$  is a *recurrent vertex*. The set of recurrent vertices is denoted by  $V_R$ .

Let  $(v_0, v_1, \dots, v_n)$  be a path from some vertex  $v_0 \in V_I$  to  $v_n$ , where  $v_i = (b_i, k_i)$  for  $i = 0, \dots, n$ . By induction, there exists  $l < |h^r(a)|$  such that  $b_i$  occurs in the image  $h^{r+iq}(x_l)$ , and in  $x$  there is a position of  $b_i$  between  $|h^{r+iq}(x_0 \cdots x_{l-1})|$  and  $|h^{r+iq}(x_0 \cdots x_l)| + 1$ , which is congruent to  $k_i$  modulo  $p$ . Note that if  $b_i$  occurs in the image  $h^{r+iq}(x_l)$ , then it occurs in the image  $h^r(x_m)$  for some letter  $x_m$  occurring in the image of  $h^{iq}(x_l)$ . The letter  $x_m$  corresponds to  $b$  occurring in (2) in the definition of the graph  $G_h$ . Note also that, for any position  $n$ , we can find  $i$  and  $l < |h^r(a)|$  such that  $x_n$  is in a position between  $|h^{r+iq}(x_0 \cdots x_{l-1})|$  and  $|h^{r+iq}(x_0 \cdots x_l)| + 1$ . Hence, we conclude that  $(b, k)$  is in  $V_R$  if and only if there exist infinitely many  $n$  such that  $x_n = b$  and  $n \equiv k \pmod{p}$ . This proves the claim.  $\square$

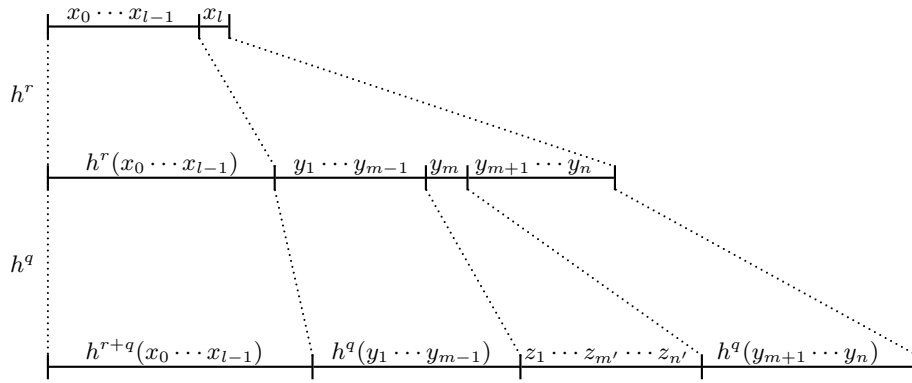


Fig. 1. Image of  $x_l$ .

The proof of the previous theorem gives us an algorithm for finding the  $k$ -sets of  $h^\omega(a)$  modulo  $p$ .

**Algorithm.** INPUT: morphism  $h$  prolongable on  $a$  and integer  $p > 0$ .

1. Find  $r$  and  $q$  satisfying (1).
2. Using  $r$  and  $q$ , construct the graph  $G_h = (V, E)$  defined in Theorem 1.
3. Find the set of recurrent vertices  $V_R$ .
4. For each letter  $b$  and integer  $k \in \{0, 1, \dots, p - 1\}$ , set  $b \in C_k$  if  $(b, k) \in V_R$ .

OUTPUT:  $k$ -sets  $C_k$  of  $h^\omega(a)$  modulo  $p$ .

## 4 Periodicity Results

In this section we show how the algorithm of the previous section can be used for solving two periodicity problems. First, let us consider a decidability question on morphic words.

**Theorem 2.** *Given a positive integer  $p$ , it is decidable whether a morphic word  $g(h^\omega(a))$  is ultimately  $p$ -periodic.*

*Proof.* Consider a morphic word  $y = (y_n)_{n \geq 0} = g(h^\omega(a))$ , where  $h: \mathcal{A}^* \rightarrow \mathcal{A}^*$  is a morphism prolongable on  $a$  and  $g: \mathcal{A}^* \rightarrow \mathcal{B}^*$  is a coding. Let  $x = (x_n)_{n \geq 0} = h^\omega(a)$ . By Theorem 1, all  $k$ -sets of  $x$  modulo  $p$  can be found algorithmically. We claim that the word  $y$  is ultimately  $p$ -periodic if and only if  $g(b) = g(c)$  for all pairs of letters  $(b, c)$  such that  $b$  and  $c$  belong to the same  $k$ -set of  $x$  modulo  $p$ .

First assume that  $g(b) = g(c)$  for all pairs  $(b, c)$  belonging to some  $k$ -set of  $x$  modulo  $p$ . Consider the subsequence  $(x_{pn+k})_{n \geq 0}$  for some  $k \in \{0, 1, \dots, p-1\}$ . After a finite prefix the subsequence contains only letters belonging to the  $k$ -set of  $x$  modulo  $p$ . Now  $g$  maps all these letters to a common letter  $b_k \in \mathcal{B}$ . Since this holds for every  $k \in \{0, 1, \dots, p-1\}$ , it follows that  $y = g(x) = uv^\omega$  where  $v = (b_0 \dots b_{p-1})^\omega$  and  $u$  is some finite word.

Conversely, assume that  $b$  and  $c$  belong to some  $k$ -set of  $x$  modulo  $p$  and  $g(b) \neq g(c)$ . Thus, there exist two increasing sequences  $(i_n)_{n \geq 1}$  and  $(j_n)_{n \geq 1}$  such that  $x_{i_n} = b$ ,  $x_{j_n} = c$  and  $i_n \equiv k \equiv j_n \pmod{p}$ . This means that, for any  $N$ , there are positions  $i, j > N$  such that  $y_i \neq y_j$  and  $i \equiv j \pmod{p}$ . Hence,  $y$  is not ultimately  $p$ -periodic.  $\square$

Another application of Theorem 1 is the following results concerning relational periods.

**Theorem 3.** *Given a positive integer  $p$ , a morphic word  $y$  and a similarity relation  $R$ , it is decidable whether  $y$  is ultimately strongly (resp. externally, weakly)  $R$ -periodic with period  $p$ .*

*Proof.* Let  $y = (y_n)_{n \geq 0} = g(h^\omega(a))$ , where  $h: \mathcal{A}^* \rightarrow \mathcal{A}^*$  is a morphism prolongable on  $a$  and  $g: \mathcal{A}^* \rightarrow \mathcal{B}^*$  is a coding. Denote  $x = (x_n)_{n \geq 0} = h^\omega(a)$ . By Theorem 1, all  $k$ -sets of  $x$  modulo  $p$  can be found algorithmically.

Similarly to the proof of the previous theorem, we can show that  $y$  is ultimately strongly  $R$ -periodic with period  $p$  if and only if  $g(b) R g(c)$  for all pairs of letters  $(b, c)$  belonging to some  $k$ -set of  $x$  modulo  $p$ .

The case of external  $R$ -periodicity is a little bit different. We claim that  $y$  is ultimately externally  $R$ -periodic with period  $p$  if and only if for each  $k \in \{0, 1, \dots, p-1\}$  there exists a letter  $b_k \in \mathcal{B}$  such that  $b_k R g(c)$  for every letter  $c$  in the  $k$ -set of  $x$  modulo  $p$ .

First assume that such letters  $b_k$  can be found. Then  $b_k$  is related to all letters occurring infinitely many times in the subsequence  $(y_{pn+k})_{n \geq 0} = (g(x_{pn+k}))_{n \geq 0}$ . Hence,  $(y_{pn+k})_{n \geq 0}$  is ultimately externally  $R$ -periodic with period 1. Since this holds for every  $k \in \{0, 1, \dots, p-1\}$ , we conclude that after a finite prefix the morphic word  $y$  is  $R$ -similar to  $(b_0 \dots b_{p-1})^\omega$ , i.e.,  $y$  is ultimately externally  $R$ -periodic with period  $p$ .

Assume next that  $y$  is ultimately externally  $R$ -periodic with period  $p$ . Then for each  $k \in \{0, 1, \dots, p-1\}$ , there exists a letter  $v_k$  such that  $y_i R v_k$  if  $i \equiv k \pmod{p}$  and  $i > N$  for some integer  $N$  large enough. Take a letter  $c$  from the  $k$ -set of  $x$  modulo  $p$ . By definition, there are infinitely many positions  $i$  such that

$x_i = c$  and  $i \equiv k \pmod{p}$ . Hence, there are infinitely many position  $i$  such that  $y_i = g(c)$  and  $i \equiv k \pmod{p}$ . By assumption, we have  $g(c) R v_k$  and the claim holds.

Finally, consider the weak  $R$ -periodicity of  $y$ . Here it suffices to find out all words of length  $p + 1$  that occur infinitely often in  $y$ . Denote the set of these factors by  $W_{p+1}$ . The word  $y$  is ultimately weakly  $R$ -periodic with period  $p$  if and only if the first and the last letter of every  $w \in W_{p+1}$  are  $R$ -similar. The set  $W_{p+1}$  can be constructed effectively. By the result of Cobham [2, Theorem 7.5.1], we may assume that  $h$  is non-erasing.

First, we find the set of all factors of  $x$  of length at most  $p + 1$ . Set

$$U_n = \bigcup_{j=1}^{p+1} F_j(h^n(a))$$

for  $n \geq 0$ . Since  $h^n(a)$  is a prefix of  $h^{n+1}(a)$  and  $x = h^\omega(a)$ , we have  $U_n \subseteq U_{n+1} \subseteq \bigcup_{j=1}^{p+1} F_j(x)$  for every  $n \geq 0$ . Since there are only finitely many factors of length at most  $p + 1$ , there must be an integer  $N$  such that  $U_N = U_{N+1}$ . Since  $h$  is non-erasing, a factor  $v$  of  $x$  of length at most  $p + 1$  is a factor of  $h(u)$  for some word  $u \in F_j(x)$  with  $j \leq p + 1$ . This implies that  $U_n = U_N$  for every  $n \geq N$  and, consequently,

$$F_{p+1} = U_N \cap \mathcal{A}^{p+1}.$$

Let  $V_{p+1}$  be the set of factors of  $x$  of length  $p + 1$  that occur infinitely often in  $x$ . To find the subset  $V_{p+1}$  of  $F_{p+1}$  we suggest a graph construction similar to that in Theorem 1. Let  $G = (V, E)$  be a directed graph, where  $V = F_{p+1}$  and there is an edge from  $u$  to  $v$  if  $v$  is a factor of  $h(u)$ . By the definition of the graph, a word  $v$  belongs to  $V_{p+1}$  if and only if there exist a vertex  $u$  such that there are infinitely long paths from  $u$  to  $v$ . The existence of such  $u$  can be easily verified. Hence, we can construct the set  $V_{p+1}$  and, clearly, we have  $W_{p+1} = g(V_{p+1})$ .  $\square$

For a relation  $R$  on the alphabet  $\mathcal{A}$ , let

$$\text{sz}(R) = |\{(a, b) \mid a R b\}|.$$

A relation  $R$  is smaller than a relation  $S$  if  $\text{sz}(R) < \text{sz}(S)$ . We can measure the degree of periodicity of an infinite word  $x \in \mathcal{A}^\omega$  by considering the smallest similarity relation  $R_{x,p}$  such that  $x$  is ultimately strongly (resp. externally, weakly)  $R_{x,p}$ -periodic with period  $p$ . It is clear that such a relation is unique. We have the following obvious corollary of Theorem 3.

**Corollary 1.** *Given a positive integer  $p$  and a morphic word  $x$ , we can effectively find the smallest similarity relation  $R_{x,p}$  such that  $x$  is ultimately strongly (resp. externally, weakly)  $R_{x,p}$ -periodic with period  $p$ .*

Unfortunately, the general ultimate periodicity problem (without specifying the period  $p$ ) for morphic words remains unsolved. Note that this very difficult and challenging problem is a special case of the *ultimate relational periodicity*

problem:

*Given a pure morphic word  $x$  and a similarity relation  $R$ , decide whether  $x$  is ultimately strongly (resp. externally, weakly)  $R$ -periodic.*

The ultimate periodicity problem for a morphic word  $g(h^\omega(a))$  is the above problem where the relation  $R$  is defined by the coding  $g$  as follows:  $a R b$  if and only if  $g(a) = g(b)$ . Note that in this case  $R$  is an equivalence relation and the definitions of strong, weak and external periods coincide.

However, using the method from [1] we can solve the ultimate relational periodicity problem restricted to automatic sequences and weak relational periods. The proof of Theorem 4 is similar to corresponding result in [1]. For completeness sake, we give the proof.

**Theorem 4.** *Given a  $k$ -automatic sequence  $x$  over  $\mathcal{A}$  and a similarity relation  $R$  on  $\mathcal{A}$ , it is decidable whether  $x$  is ultimately weakly  $R$ -periodic.*

*Proof.* Let  $x = (x_n)_{n \geq 0}$  be an infinite  $k$ -automatic sequence. The sequence  $x$  is ultimately weakly  $R$ -periodic if there exists  $P \geq 1$ ,  $N \geq 0$  such that  $x_i R x_{i+P}$  for all  $i \geq N$ .

First, construct a nondeterministic finite automaton (NFA) that on input  $(P, N)$  successively “guesses” the base- $k$  digits of  $I$  starting with the least significant digit and accepts if  $I \geq N$  and  $(x_I, x_{I+P}) \notin R$ . Let  $M = (Q, q_0, \Sigma_k, \delta, \mathcal{B}, \tau)$  be the finite automaton with output that generates the sequence  $x$ , i.e.,  $x_n = \tau(\delta(q_0, \text{rep}_k(n)))$  for all  $n \geq 0$ . Then define  $M_1 = (Q', q'_0, \Sigma_k \times \Sigma_k, \delta', F')$ , where

$$\begin{aligned} Q' &= \{<, =, >\} \times \{0, 1\} \times Q \times Q, \\ q'_0 &= [=, 0, q_0, q_0], \\ F' &= \{[b, 0, q, r] \mid b \in \{>, =\} \text{ and } (\tau(q), \tau(r)) \notin R\}. \end{aligned}$$

Here  $F'$  is the set of accepting states, and the input is actually a sequence of pairs

$$(p_0, n_0)(p_1, n_1)(p_2, n_2) \cdots (p_j, n_j),$$

where  $p_j p_{j-1} \cdots p_0$  is a base- $k$  representation of  $P$  and  $n_j n_{j-1} \cdots n_0$  is a base- $k$  representation of  $N$ , either one or both padded with leading zeros to ensure that their lengths are equal.

In order to verify that  $I \geq N$ , we have a flag  $\{<, =, >\}$  that tells us how the digits of the guessed  $I$  seen so far are in relation to the digits of the input  $N$  read so far. Let  $i'$  be the next guessed digit of  $I$  and let  $n'$  be the next read digit of  $N$ . We update the flag as follows:

$$\begin{aligned} u(<, i', n') &= \begin{cases} <, & \text{if } i' \leq n'; \\ >, & \text{if } i' > n'; \end{cases} \\ u(=, i', n') &= \begin{cases} <, & \text{if } i' < n'; \\ =, & \text{if } i' = n'; \\ >, & \text{if } i' > n'; \end{cases} \\ u(>, i', n') &= \begin{cases} <, & \text{if } i' < n'; \\ >, & \text{if } i' \geq n'. \end{cases} \end{aligned}$$



In a state  $[b, c, q, r] \in Q'$  the component  $b$  is the flag,  $c$  is the carry bit in the computation  $I + P$ ,  $q$  is the state in  $M$  reached by the digits of  $I$  guessed so far, and  $r$  is the state in  $M$  reached by the digits of  $I + P$  calculated so far. Hence, the transitions  $\delta': Q' \times (\Sigma_k \times \Sigma_k) \rightarrow 2^{Q'}$  are defined by

$$\delta'([b, c, q, r](p', n')) = \left\{ \left[ u(b, i', n'), \left\lfloor \frac{i' + p' + c}{k} \right\rfloor, \delta(q, i'), \delta(r, (i' + p' + c) \bmod k) \right] \mid 0 \leq i' < k \right\}.$$

We must ensure that the acceptance of  $(P, N)$  does not depend on the number of leading zeros and that the calculation is correct in the case where the guessed  $I$  is longer than the input. Hence, we modify the accepting states of  $M_1$  by constructing a new NFA  $\hat{M}_1 = (Q', q'_0, \Sigma_k \times \Sigma_k, \delta', \hat{F})$  where a state  $[b, c, q, r]$  belongs to  $\hat{F}$  if there exists  $j \geq 0$  such that  $\delta'([b, c, q, r], (0, 0)^j)$  contains a state in  $F'$ .

Then convert  $\hat{M}_1$  to a deterministic finite automaton (DFA)  $M_2$  using the subset construction. By interchanging accepting and non-accepting states we obtain a DFA  $M_3$  that accepts  $(P, N)$  if and only if  $x_I R x_{I+P}$  for all  $I \geq N$ . Now  $x$  is ultimately weakly  $R$ -periodic if and only if  $M_3$  accepts some input  $(P, N)$  with  $P \geq 1$ . This can be checked by creating a DFA  $M_4$  that accepts  $\Sigma_k^*(\Sigma_k \setminus \{0\})\Sigma_k^* \times \Sigma_k^*$  and forming the direct product DFA  $M_5$  that accepts exactly the words accepted by both  $M_3$  and  $M_4$ . Thus, the word  $x$  is ultimately weakly  $R$ -periodic if and only if the language accepted by  $M_5$  is not empty. This can be easily checked.  $\square$

## References

1. Allouche, J.-P., Rampersad, N., Shallit, J.: Periodicity, repetitions, and orbits of an automatic sequence. *Theoret. Comput. Sci* 410, 2795–2803 (2009)
2. Allouche, J.-P., Shallit, J.: Automatic sequences: Theory, Applications, Generalizations. Cambridge University Press (2003)
3. Césari, Y., Vincent, M.: Une caractérisation des fonctions périodiques. *C. R. Acad. Sci. Paris* 286(A), 1175–1177 (1978)
4. Cobham, A.: Uniform tag sequences. *Math. Systems Theory* 6, 164–192 (1972)
5. Duval, J.-P.: Périodes et répétitions des mots du monoïde libre. *Theoret. Comput. Sci.* 9, 17–26 (1979)
6. Fine, N.J., Wilf, H.S.: Uniqueness theorem for periodic functions. *Proc. Amer. Math. Soc.* 16, 109–114 (1965)
7. Halava, V., Harju, T., Kärki, T.: Interaction properties of relational periods. *Discrete Math. Theor. Comput. Sci.* 10, 87–112 (2008)
8. Harju, T., Linna, M.: On the periodicity of morphisms on free monoids. *Theoret. Inform. Appl.* 20, 47–54 (1986)
9. Hedlund, G.A., Morse, M.: Symbolic dynamics II: Sturmian trajectories. *Amer. J. Math.* 62, 1–42 (1940)
10. Honkala, J.: A decision method for the recognizability of sets defined by number systems. *Theoret. Inform. Appl.* 20, 395–403 (1986)

11. Mignosi, F., Restivo, A.: Periodicity. In: Lothaire, M. (ed.) Algebraic Combinatorics on Words, Encyclopedia of Mathematics and its Applications, 90. Cambridge University Press, Cambridge, 269–311 (2002)
12. Pansiot, J.-J.: Decidability of periodicity for infinite words. Theoret. Inform. Appl. 20, 43–46 (1986)