RAIRO-Theor. Inf. Appl. 49 (2015) 61–65
DOI: 10.1051/ita/2014028

Available online at:
www.rairo-ita.org

THE FINITENESS PROBLEM FOR MONOIDS OF MORPHISMS

Juha Honkala[1]

Abstract. We study finitely generated monoids consisting of endomorphisms of a free monoid. We give a necessary and sufficient condition for such a monoid to be infinite and show that this condition is decidable. As a special case we discuss the morphism torsion problem.

Mathematics Subject Classification. 20M05, 68Q45.

1. Introduction and results

Let X be a finite alphabet and let $\text{Hom}(X^*)$ be the monoid consisting of all endomorphisms of the free monoid X^*. The problem whether a morphism $h \in \text{Hom}(X^*)$ generates a finite submonoid of $\text{Hom}(X^*)$ is called the *morphism torsion problem* in [2], where it is shown that this problem is decidable in polynomial time. The idea of the proof by Cassaigne and Nicolas is to replace the morphism h by its incidence matrix M and check whether M generates a finite matrix monoid. This idea of Cassaigne and Nicolas can also be used to decide whether a finite subset of $\text{Hom}(X^*)$ generates a finite monoid by applying the decidability of the finiteness problem for finitely generated matrix monoids due to [3, 5]. The decidability of the finiteness problem for finitely generated morphism monoids is perhaps not stated explicitly in the literature but it follows, for example, as a special case of the decidability of the finiteness problem for ET0L languages (see [4]). However, decision methods based on [3, 5] or the properties of ET0L languages lead to algorithms with very high complexity.

In this note we give a different approach to these problems. Our approach is largely inspired by a classical result of Salomaa characterizing exponential D0L

Keywords and phrases. Free monoid morphism, finiteness problem, decidability.

[1] Department of Mathematics and Statistics, University of Turku, 20014 Turku, Finland. juha.honkala@utu.fi

systems (see [7]). By definition, a *D0L system* is a triple $G = (X, h, w)$, where X is a finite alphabet, h is a morphism in $\mathrm{Hom}(X^*)$ and w is a word in X^*. G is called *exponential* if there exists a real number $\alpha > 1$ such that the length of $h^n(w)$ exceeds α^n for all large values of n. Salomaa proved that $G = (X, h, w)$ is exponential if and only if there is a letter $x \in X$ occurring in some $h^n(w)$ and an integer k such that the word $h^k(x)$ contains at least two occurrences of x.

Our characterization of the finite subsets of $\mathrm{Hom}(X^*)$ generating an infinite submonoid of $\mathrm{Hom}(X^*)$ uses the notion of a cyclic letter. If $x \in X$ and $w \in X^*$, then $|w|_x$ is the number of occurrences of x in w. More generally, if $Y \subseteq X$, then $|w|_Y$ is the number of occurrences of the letters of Y in w. Now, if $h \in \mathrm{Hom}(X^*)$, the set $\mathrm{CYCLIC}(h)$ consisting of *cyclic letters with respect to* h is defined by

$$\mathrm{CYCLIC}(h) = \{x \in X \mid |h(x)|_x \geq 1\}.$$

Using this notion we can state our main result.

Theorem 1.1. *Let X be a finite alphabet and let $h_1, \ldots, h_m \in Hom(X^*)$. Then $h_1, \ldots, h_m$ generate an infinite submonoid of $Hom(X^*)$ if and only if there exists a morphism h in the submonoid generated by $h_1, \ldots, h_m$ and a letter $x \in CYCLIC(h)$ such that*

$$|h(x)|_{\mathrm{CYCLIC}(h)} \geq 2. \tag{1.1}$$

Theorem 1.1 will be proved in Section 3. In Section 4 we will show that the condition of Theorem 1.1 is decidable. In Section 2 we will discuss the morphism torsion problem.

2. The morphism torsion problem

Theorem 1.1 takes a very simple form if we consider monoids generated by a single morphism. We will use the well-known fact that if X is a d-letter alphabet, $w \in X^*$ and $h \in \mathrm{Hom}(X^*)$, then $h^n(w) \neq \varepsilon$ for all $n \geq 0$ if and only if $h^d(w) \neq \varepsilon$. To prove this fact it suffices to show that if $h^n(x) = \varepsilon$ for some $n \geq 0$ and $x \in X$, then $h^d(x) = \varepsilon$. We prove this inductively. If $d = 1$, the claim holds. Assume $d > 1$ and assume that the claim holds for alphabets which have less than d letters. Now, x is not a factor of $h^m(x)$ for any $m \geq 1$. Hence $h(x)$ is a word over the $(d-1)$-letter alphabet $X - \{x\}$ and h defines an endomorphism of $(X - \{x\})^*$. The claim follows inductively.

Theorem 2.1. *Let X be an alphabet having d letters and let $h \in Hom(X^*)$. Then h generates an infinite submonoid of $Hom(X^*)$ if and only if there is a positive integer $k \leq d$, a letter $x \in X$ and words $u, v \in X^*$ such that*

$$h^k(x) = uxv \tag{2.1}$$

and

$$h^d(uv) \neq \varepsilon. \tag{2.2}$$

Proof. Suppose first that there is a positive integer k, a letter $x \in X$ and words $u, v \in X^*$ such that (2.1) and (2.2) hold. Then for $n \geq 1$,

$$h^{nk}(x) = h^{(n-1)k}(u)h^{(n-2)k}(u) \dots h^k(u)uxvh^k(v) \dots h^{(n-2)k}(v)h^{(n-1)k}(v).$$

Because $h^i(uv) \neq \varepsilon$ for all i, it follows that h generates an infinite submonoid of $\mathrm{Hom}(X^*)$.

Conversely, assume that h generates an infinite submonoid of $\mathrm{Hom}(X^*)$. Choose recursively letters $x_0, x_1, \dots, x_d \in X$ such that $\{h^n(x_i) \mid n \geq 0\}$ is infinite for $i = 0, 1, \dots, d$ and x_{i+1} is a factor of $h(x_i)$ for $i = 0, 1, \dots, d-1$. Then there exist integers s and t such that $s < t$ and $x_s = x_t$. Hence x_s is a factor of $h^{t-s}(x_s)$. Let $u, v \in X^*$ be words such that

$$h^{t-s}(x_s) = ux_sv.$$

Then $h^d(uv) \neq \varepsilon$. Indeed, if $h^d(uv) = \varepsilon$, the set $\{h^{n(t-s)}(x_s) \mid n \geq 0\}$ would be finite which is not possible because $\{h^n(x_s) \mid n \geq 0\}$ is infinite. □

Observe that the criterion of Theorem 2.1 leads to a polynomial time algorithm.

3. Proof of Theorem 1.1

We first recall some facts concerning incidence matrices of morphisms. Let the letters of X be $x_1, \dots, x_d$. Then the *incidence matrix* M_h of $h \in \mathrm{Hom}(X^*)$ is defined by

$$M_h = \begin{pmatrix} |h(x_1)|_{x_1} & |h(x_2)|_{x_1} & \dots & |h(x_d)|_{x_1} \\ |h(x_1)|_{x_2} & |h(x_2)|_{x_2} & \dots & |h(x_d)|_{x_2} \\ \vdots & \vdots & \ddots & \dots \\ |h(x_1)|_{x_d} & |h(x_2)|_{x_d} & \dots & |h(x_d)|_{x_d} \end{pmatrix}.$$

If $g, h \in \mathrm{Hom}(X^*)$ we have $M_gM_h = M_{gh}$. If M is a $d \times d$ matrix with nonnegative integer entries, there exist at most finitely many $h \in \mathrm{Hom}(X^*)$ such that $M = M_h$.

Let now X be a finite alphabet having d letters and let $h_1, \dots, h_m \in \mathrm{Hom}(X^*)$. Let H be the submonoid of $\mathrm{Hom}(X^*)$ generated by $h_1, \dots, h_m$. If $h \in \mathrm{Hom}(X^*)$ is an arbitrary morphism, let h^* be the submonoid of $\mathrm{Hom}(X^*)$ generated by h. In other words,

$$H = \{h_{i_1}h_{i_2} \dots h_{i_p} \mid p \geq 0, i_1, \dots, i_p \in \{1, \dots, m\}\}$$

and

$$h^* = \{h^i \mid i \in \mathbb{N}\}.$$

In the proof of the following lemma we use the result of McNaughton and Zalcstein (see [1,6]) stating that if every matrix of a finitely generated monoid $\mathcal{M}$ of matrices over $\mathbb{Q}$ generates a finite monoid, then the monoid $\mathcal{M}$ is finite.

Lemma 3.1. *Using the notation explained above, H is a finite monoid if and only if h^* is a finite monoid for all $h \in H$.*

Proof. If H is finite, so is h^* for all $h \in H$. Conversely, suppose h^* is a finite monoid for all $h \in H$. Then M_h generates a finite monoid for all $h \in H$. Hence, the matrix monoid $\mathcal{M}$ generated by $M_{h_1}, \dots, M_{h_m}$ has the property that every matrix of $\mathcal{M}$ generates a finite submonoid of $\mathcal{M}$. Hence by the result of McNaughton and Zalcstein quoted above, $\mathcal{M}$ is a finite monoid. Therefore also H is finite. □

Now we are ready to prove Theorem 1.1.

Proof of Theorem 1.1. Suppose first that there is a morphism $h \in H$ and a letter $x \in \mathrm{CYCLIC}(h)$ such that (1.1) holds. Then

$$|h^n(x)|_{\mathrm{CYCLIC}(h)} \geq n + 1 \tag{3.1}$$

for all $n \geq 1$. Indeed, (3.1) holds by assumption if $n = 1$. If (3.1) holds for $n \geq 1$, then $h^n(x)$ contains x and at least n other occurrences of letters which are cyclic with respect to h. Hence $h^{n+1}(x)$ contains at least $n + 2$ cyclic letters, because h produces at least one cyclic letter when applied to any cyclic letter and h produces at least two cyclic letters when applied to x.

Now (3.1) implies that h^* and H are infinite.

Suppose then that H is infinite. By Lemma 3.1 there exists $h \in H$ such that h^* is infinite. By Theorem 2.1 there is a positive integer k, a letter $x \in X$ and words $u, v \in X^*$ such that (2.1) and (2.2) hold. Let $g = h^k$. Then $g(x) = uxv$ and $g^n(uv) \neq \varepsilon$ for all $n \geq 0$. If there is a positive integer n such that $g^n(x)$ contains at least two occurrences of x the condition of Theorem 1.1 holds. Suppose that

$$|g^n(x)|_x = 1$$

for all $n \in \mathbb{N}$. Choose recursively letters $y_1, \dots, y_d \in X - \{x\}$ such that $g^n(y_i) \neq \varepsilon$ for all $n \in \mathbb{N}$ and $i = 1, \dots, d$ and y_{i+1} is a factor of $g(y_i)$ for $i = 1, \dots, d-1$ and y_1 is a factor of $g(x)$. Then there exist integers p and q such that $p < q$ and $y_p = y_q$. Hence $y_p, \dots, y_{q-1} \in \mathrm{CYCLIC}(g^{q-p})$. Now, choose an integer j such that $p \leq j < q$ and $q - p$ divides j. Let $f = g^j$. Then $f \in H$ and

$$|f(x)|_{\mathrm{CYCLIC}(f)} \geq 2.$$

Indeed, $f(x)$ contains the letters x and y_j which are cyclic with respect to f and $x \neq y_j$. This shows that the condition of Theorem 1.1 holds. □

4. Decidability

Let $X = \{x_1, \dots, x_d\}$ be an alphabet having d letters and let $h_1, \dots, h_m \in \mathrm{Hom}(X^*)$. Let H be the submonoid of $\mathrm{Hom}(X^*)$ generated by $h_1, \dots, h_m$. In this section we show that the condition of Theorem 1.1 is decidable. In other words, we show that it is decidable whether there exist $h \in H$ and $x \in \mathrm{CYCLIC}(h)$ such that (1.1) holds.

First, if $h \in \mathrm{Hom}(X^*)$, define the mapping $A(h) : X \to \mathbb{N}^d$ by

$$A(h)(x) = (v_1, \ldots, v_d)$$

where

$$v_i = \begin{cases} 2 & \text{if } |h(x)|_{x_i} \geq 2 \\ 1 & \text{if } |h(x)|_{x_i} = 1 \\ 0 & \text{if } |h(x)|_{x_i} = 0 \end{cases}$$

for $i = 1, \ldots, d$. Now, if $f, g, h \in \mathrm{Hom}(X^*)$, then $A(f) = A(g)$ implies that $A(hf) = A(hg)$. From this it follows that we can compute the finite set

$$\mathcal{A} = \{A(h) \mid h \in H\}.$$

The decidability of the condition of Theorem 1.1 follows. Indeed, there exist $h \in H$ and $x \in \mathrm{CYCLIC}(h)$ such that (1.1) holds if and only if there exists an integer $i \in \{1, \ldots, d\}$ and $\alpha \in \mathcal{A}$ such that the ith coordinate of $\alpha(x_i)$ is two or there exist $i, j \in \{1, \ldots, d\}$, $i \neq j$ such that the ith coordinate of $\alpha(x_i)$ equals 1 and the jth coordinates of $\alpha(x_i)$ and $\alpha(x_j)$ are 1 or 2.

Hence Theorem 1.1 gives a simple algorithm for the finiteness problem for monoids of morphisms.

Acknowledgements. The author would like to thank the referees for useful suggestions.

References

[1] J. Berstel and C. Reutenauer, Noncommutative Rational Series with Applications. Cambridge University Press (2011).

[2] J. Cassaigne and F. Nicolas, On the decidability of semigroup freeness. *RAIRO: ITA* **46** (2012) 355–399.

[3] G. Jacob, La finitude des représentations linéaires des semi-groupes est décidable. *J. Algebra* **52** (1978) 437–459.

[4] L. Kari, G. Rozenberg and A. Salomaa, L systems. In vol. 1 of *Handbook of Formal Languages*, edited by G. Rozenberg and A. Salomaa. Springer (1997) 253–328.

[5] A. Mandel and I. Simon, On finite semigroups of matrices. *Theoret. Comput. Sci.* **5** (1977) 101–111.

[6] R. McNaughton and Y. Zalcstein, The Burnside problem for semigroups. *J. Algebra* **34** (1975) 292–299.

[7] A. Salomaa, On exponential growth in Lindenmayer systems. *Indag. Math.* **35** (1973) 23–30.

Communicated by P. Weil.

Received April 22, 2014. Accepted November 5, 2014.