INFORMATIQUE THÉORIQUE ET APPLICATIONS

T. HARJU
L. ILIE

Languages obtained from infinite words

Informatique théorique et applications, tome 31, nº 5 (1997), p. 445-455
<http://www.numdam.org/item?id=ITA_1997__31_5_445_0>

Informatique théorique et Applications/Theoretical Informatics and Applications
(vol. 31, n° 5, 1997, pp. 445-455)

LANGUAGES OBTAINED FROM INFINITE WORDS (*)

by T. Harju (1) and L. Ilie (**, 2, †)

Abstract. – *We prove that it is decidable whether or not a regular language can be written as the set of all finite factors of an infinite word. The result holds for both right-infinite and bi-infinite words.*

Résumé. – *Nous démontrons qu'il est indécidable de savoir si un language rationnel peut être décrit comme l'ensemble des facteurs d'un mot infini. Le résultat vaut aussi bien pour les mots infinis à droite que pour les mots infinis à gauche et à droite.*

1. INTRODUCTION AND BASIC DEFINITIONS

There are several classical ways to associate a set of finite words to an infinite word α. One can take the set of all finite prefixes or finite factors of α, *Pref* (α), or *Fact* (α), respectively, [MaPa1], [MaPa2], or the set of all finite words which are not prefixes of α, *Copref* (α), [AuGa], [AFG], [Ber]. Conversely, for a language of finite words, one can associate infinite words considering the notions of limit [Ei] or adherence [BoNi].

This paper is devoted to the study of languages obtained from infinite words by taking the set of all finite factors. More precisely, we prove that it is decidable whether or not a regular language can be written as the set of all finite factors of a right-infinite word, answering an open problem in [MaPa1], [MaPa2], and, then, we show that the same holds for bi-infinite words. We mention that [Bea1], [Bea2], and [BeaN] deal with related problems.

For an alphabet Σ, we denote by Σ^* the set of all finite words over Σ and by Σ^ω the set of all (one-sided) infinite words over Σ; λ denotes the empty word and $\Sigma^+ = \Sigma^* - \{\lambda\}$. For a finite word $w \in \Sigma^*$, we denote by

(*) Received August 30, 1997.
(**) Research supported by the Academy of Finland, Project 137358.
(1) Mathematics Department, University of Turku, FIN-20014 Turku, Finland.
(2) Turku Centre for Computer Science TUCS, FIN-20520 Turku, Finland.
(†) On leave of absence from Faculty of Mathematics, University of Bucharest, Str. Academiei 14, R-70109, Bucharest, Romania.

$|w|$ the length of w. For a finite (non-empty) word $w \in \Sigma^+$, we denote by w^ω the infinite word $w^\omega = www\ldots$ For all formal language theory notions and results we refer to [HoUl] and [Sa].

For an infinite word $\alpha \in \Sigma^\omega$, we denote by *Fact* (α) the set of all finite factors of α. For a language $L \subseteq \Sigma^*$, *Fact* (L) is the set of all factors of words in L. Also, we denote by $\mathcal{F}_{fact}$ the family of languages of the form *Fact* (α), for an arbitrary infinite word α, that is,

$$\mathcal{F}_{fact} = \{L |\ \text{there are } \Sigma \text{ and } \alpha \in \Sigma^\omega \text{ such that } L = Fact\,(\alpha)\}.$$

2. FACTORS OF INFINITE WORDS

In [MaPa1] it is proved that it is undecidable whether an arbitrarily given context-free language is in the family $\mathcal{F}_{fact}$ or not. The same problem for regular languages is left open. In the following, we solve this problem in the affirmative.

For a regular language $R \subseteq \Sigma^*$, a finite automaton $\mathcal{A} = (Q, \Sigma, \delta, I, F)$ recognizing R and having a deterministic transition function $\delta : Q \times \Sigma \to Q$ is called *strongly minimal* if and only if no state or transition in $\mathcal{A}$ is useless or redundant. That is, if we eliminate any state or transition in $\mathcal{A}$, the obtained automaton recognizes a language strictly contained in R.

LEMMA 2.1: *Any regular language $R \subseteq \Sigma^*$ closed under taking factors is recognized by a strongly minimal automaton $\mathcal{A} = (Q, \Sigma, \delta, Q, Q)$ in which all states are both initial and final.*

Proof: Let $\mathcal{A}' = (Q', \Sigma, \delta', I, T)$ be the minimal deterministic finite automaton recognizing R. Consider the automaton $\mathcal{A}'' = (Q', \Sigma, \delta', Q', Q')$. Since R closed under taking factors, $R = L(\mathcal{A}'')$. Because the equivalence problem is decidable for finite automata, we can now iteratively eliminate from $\mathcal{A}''$ all states and transitions which are either useless or redundant. That is, if $s \in Q'$ (or $\delta'(s, a) = s'$, for $s, s' \in Q$, $a \in \Sigma$) and the language recognized by the finite automaton $\mathcal{B}$ obtained from $\mathcal{A}''$ by removing the state s together with all transitions containing it (or removing the transition by a from s to s', respectively) is R, then take $\mathcal{B}$ instead of $\mathcal{A}''$ and continue the reduction. Obviously, after a finite number of steps, the automaton $\mathcal{A}$ asked for in the claim of our lemma is obtained. (Note that $\mathcal{A}$ is not necessarily unique, but this will not cause troubles later.) ■

For a finite automaton $\mathcal{A} = (Q, \Sigma, \delta, I, F)$, denote by $G(\mathcal{A})$ the graph associated to $\mathcal{A}$ and define the relation $\rightarrow \subseteq Q \times Q$ by

$$p \rightarrow q \text{ if and only if there is a path from } p \text{ to } q \text{ in } G(\mathcal{A}).$$

The relation $\equiv \subseteq Q \times Q$ defined by $p \equiv q$ if and only if $p \rightarrow q$ and $q \rightarrow p$ is an equivalence relation which induces an acyclic structure on $Q/_{\equiv}$, *i.e.*, the graph $G = (Q/_{\equiv}, E)$ with $Q/_{\equiv}$ as the set of vertices and with the set of edges

$$E = \{([p], [q]) | [p], [q] \in Q/_{\equiv} \text{ and } p \rightarrow q\}$$

is acyclic. (We have denoted the equivalence class of $p \in Q$ with respect to $\equiv$ by $[p]$.)

The automaton $\mathcal{A}$ is called *diconnected* if and only if there is a state $q \in Q$ such that $Q = [q]$. (For instance, the restriction of any automaton to an equivalence class with respect to the relation $\equiv$ is a diconnected automaton.)

An equivalence class $[p] \in Q/_{\equiv}$ is called *trivial* if and only if $[p]$ is a singleton ($[p] = \{p\}$) and there is no transition $p \xrightarrow{a} p$, for any $a \in \Sigma$.

A finite automaton $\mathcal{A} = (Q, \Sigma, \delta, I, F)$ is called *ultimately periodic* if and only if there is a state $q \in Q$ such that

$$\begin{aligned} Q - [q] &= \{s_1, s_2, \ldots, s_k\}, && \text{for some } k \geq 1, \\ [q] &= \{q_1 = q, q_2, \ldots, q_l\}, && \text{for some } l \geq 1, \end{aligned}$$

and all transitions in $\mathcal{A}$ are

$$\begin{aligned} &\delta(s_i, a_i) = s_{i+1}, 1 \leq i \leq k-1, && \text{for some } a_1, a_2, \ldots, a_{k-1} \in \Sigma, \\ &\delta(s_k, a_k) = q, && \text{for some } a_k \in \Sigma, \\ &\delta(q_j, b_j) = q_{j+1}, 1 \leq j \leq l-1, && \text{for some } b_1, b_2, \ldots, b_{l-1} \in \Sigma, \\ &\delta(q_l, b_l) = q, && \text{for some } b_l \in \Sigma. \end{aligned}$$

Informally speaking, a finite automaton $\mathcal{A}$ is ultimately periodic if and only if $G(\mathcal{A})$ has the form in Figure 1 (using the notations in the definition).

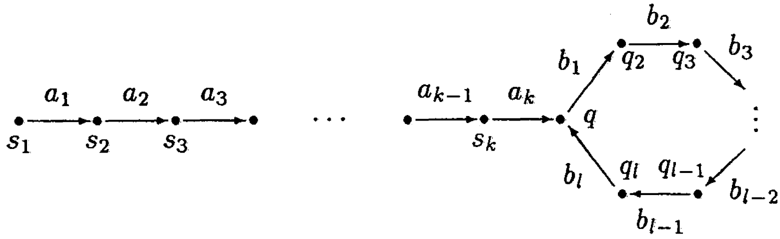


Figure 1.

(Note that if $\mathcal{A}$ is ultimately periodic, then $\mathcal{A}$ is not diconnected. Moreover, for any $p \in Q - [q]$, $[p]$ is trivial.)

An equivalence class $[p] \in Q/_{\equiv}$ is called a *source* if and only if there is no $[q] \in Q/_{\equiv} - [p]$ such that $q \to p$.

LEMMA 2.2: *For a regular language $R \subseteq \Sigma^*$ such that $R \in \mathcal{F}_{fact}$, let $\mathcal{A} = (Q, \Sigma, \delta, Q, Q)$ be a strongly minimal automaton constructed using Lemma 2.1 for R. If $\mathcal{A}$ is not diconnected, then there is a unique equivalence class $[p] \in Q/_{\equiv}$ which is a source. Moreover, $[p]$ is trivial and there is exactly one transition leaving p in $\mathcal{A}$.*

Proof: The existence of a source is guaranteed by the fact that the graph $G = (Q/_{\equiv}, E)$ is acyclic.

Let us show that any source is trivial. For, take a source $[p] \in Q/_{\equiv}$ and consider the subautomata

$$\mathcal{A}^{[p]} = ([p], \Sigma, \delta|_{[p]}, [p], [p])$$

and

$$\mathcal{A}^{Q-[p]} = (Q - [p], \Sigma, \delta|_{Q-[p]}, Q - [p], Q - [p])$$

of $\mathcal{A}$.

Suppose, contrary to the claim, that $[p]$ is not trivial. Take $p_1, p_2 \in [p]$ such that if $card([p]) \geq 2$, then $p_1 \neq p_2$, otherwise $p_1 = p_2 = p$. If $p_1 \neq p_2$, then, as $\mathcal{A}^{[p]}$ is diconnected, there must be in $\mathcal{A}^{[p]}$ a path from p_1 to p_2, say $p_1 \xrightarrow{y} p_2$, $y \in \Sigma^+$, and another one from p_2 to p_1, say $p_2 \xrightarrow{z} p_1$, $z \in \Sigma^+$. If $p_1 = p_2$, as $[p]$ is not trivial, there must be a transition $p \xrightarrow{a} p$, for some $a \in \Sigma$, and we can take $y = z = a$. So, in what follows, it will not be important whether $p_1 \neq p_2$ or not. What is important is the fact that, in both cases, the word yz is not empty.

As, clearly, $G(\mathcal{A})$ is connected and the sets of states of $\mathcal{A}^{[p]}$ and $\mathcal{A}^{Q-[p]}$, respectively, are non-empty (p is in $\mathcal{A}^{[p]}$ and, if $\mathcal{A}^{Q-[p]}$ is empty, then $\mathcal{A} = \mathcal{A}^{[p]}$ hence diconnected, a contradiction), there must be a path from a state in $\mathcal{A}^{[p]}$ to one in $\mathcal{A}^{Q-[p]}$ and we can find $p_3 \in [p]$, $q_1 \in Q - [p]$, and a transition $p_3 \xrightarrow{a} q_1$, $a \in \Sigma$, in $\mathcal{A}$. Since $\mathcal{A}$ is strongly minimal, there is a word $w \in \Sigma^*$ which contains a and is accepted by the automaton $\mathcal{A}$ but not accepted by the automaton obtained from $\mathcal{A}$ by removing the transition $p_3 \xrightarrow{a} q_1$. (That is, when $\mathcal{A}$ accepts w, then it must read a from p_3 to q_1.) It follows that we can find the states $p_4 \in [p]$, $q_2 \in Q - [p]$ and the words $w_1, w_2 \in \Sigma^*$ such that $w = w_1 a w_2$ and there are paths $p_4 \xrightarrow{w_1} p_3$ in $\mathcal{A}^{[p]}$ and $q_1 \xrightarrow{w_2} q_2$ in $\mathcal{A}^{Q-[p]}$. Using again the fact that $\mathcal{A}^{[p]}$ is diconnected, we get a path from p_1 to p_4 in $\mathcal{A}^{[p]}$, say $p_1 \xrightarrow{u} p_4$, $u \in \Sigma^*$, *see* Figure 2 below.

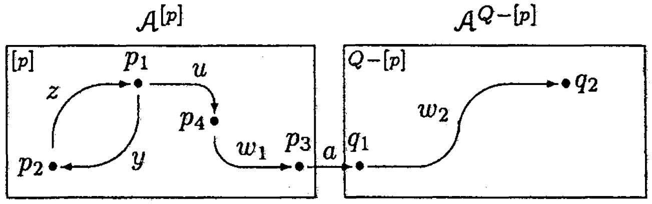
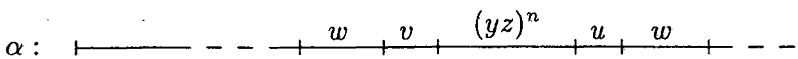


Figure 2.

Then $(yz)^* u w_1 a w_2 \subseteq R$. Assume $R = Fact(\alpha)$ for some infinite word $\alpha \in \Sigma^\omega$ (there is such an α by hypothesis). As the language $(yz)^*$ contains arbitrarily long words, there must be an $n \geq 0$ such that an occurence of $(yz)^n uw$ appears in α after an occurrence of w ("after" meaning at a larger distance from the beginning of α). Supposing that we have



for some $v \in \Sigma^*$, α has a factor $wv(yz)^n uw$. But now $wv(yz)^n uw \in R$. Because, when accepting w, $\mathcal{A}$ must read a from p_3 to q_1, using the fact that $[p]$ is a source, we obtain that the word $v(yz)^n uw$ is accepted by $\mathcal{A}$ without reading $p_3 \xrightarrow{a} q_1$. In particular, w is accepted by $\mathcal{A}$ without reading $p_3 \xrightarrow{a} q_1$, a contradiction. If follows that $[p]$ is trivial.

Let us prove now that there is exactly one source in $\mathcal{A}$. For this, suppose that $[p]$ and $[q]$ are two different sources. As shown above, $[p] = \{p\}$ and $[q] = \{q\}$. As $\mathcal{A}$ is strongly minimal, there must be transitions starting from p and q labeled by $a \in \Sigma$ and $b \in \Sigma$, respectively. Moreover, there are some words ax_1, $bx_2 \in L(\mathcal{A})$ such that $\mathcal{A}$ must read the transition which leaves $p(q)$ and is labeled by $a(b)$ in order to accept the word ax_1 (bx_2, respectively). It follows that the word ax_1 cannot be prolonged to the left in R, that is, there is no word $w \in \Sigma^+$ such that $wax_1 \in R$. As $R = Fact(\alpha)$, $\alpha \in \Sigma^\omega$, α must begin with ax_1. Since the same resoning can be done for bx_2, we get that of ax_1 and bx_2 one is a prefix of the other. Suppose that ax_1 is a prefix of bx_2. In this case, ax_1 is accepted by the automaton $\mathcal{A}$ without reading the transition labeled a leaving p, a contradiction.

A similar reasoning shows that for a source $[p]$ the number of transitions leaving p is exactly one. Indeed, it is at least one and if there are two transitions labeled a and b, then $a \neq b$, because $\mathcal{A}$ is deterministic, and then α must start with both a and b, a contradiction. ■

LEMMA 2.3: *For an infinite regular language $R \subseteq \Sigma^*$, $R \in \mathcal{F}_{fact}$ if and only if an automaton $\mathcal{A}$ constructed for R in Lemma 2.1 is either diconnected or ultimately periodic.*

Proof: Suppose first that $\mathcal{A} = (Q, \Sigma, \delta, Q, Q)$ is diconnected and take an arbitrary $q \in Q$. As R is infinite, there must be at least one cycle $q \to q$ in $G(\mathcal{A})$ hence there are infinitely many such cycles, the set of them being

$$C(q) = \{w \in \Sigma^* | q \xrightarrow{w} q\} = \{w_1, w_2, w_3 \ldots\}.$$

Define

$$\alpha = w_1 w_2 w_3 \ldots$$

We have $L(\mathcal{A}) = Fact(\alpha)$. The inclusion $Fact(\alpha) \subseteq L(\mathcal{A})$ is trivial. To prove the other one, take $w \in L(\mathcal{A})$. There are p_1, $p_2 \in Q$ such that $p_1 \xrightarrow{w} p_2$. Since $\mathcal{A}$ is diconnected, we have also $q \xrightarrow{u} p_1$, $p_2 \xrightarrow{v} q$, for some u, $v \in \Sigma^*$. If follows that there is an $i \geq 1$ such that $w_i = uwv$, so $w \in Fact(\alpha)$.

If $\mathcal{A}$ is ultimately periodic then $G(\mathcal{A})$ has the form in Figure 1. Hence (using the notations there)

$$L(\mathcal{A}) = Fact(a_1 a_2 \ldots a_k (b_1 b_2 \ldots b_l)^\omega).$$

Conversely, take a regular language $R \subseteq \Sigma^*$ such that $R \in \mathcal{F}_{Fact}$ and construct $\mathcal{A}$ as in Lemma 2.1. If $\mathcal{A}$ is diconnected, then we are done. Otherwise, by Lemma 2.2, there is exactly one source, say $[p_1]$, in $\mathcal{A}$. Morevoer, $[p_1]$ is trivial and there is exactly one transition in $\mathcal{A}$ starting from p_1, say $p_1 \xrightarrow{a_1} p_2$, for some $p_2 \in Q - \{p_1\}$ and $a_1 \in \Sigma$. If $\alpha \in \Sigma^\omega$ with $R = Fact(\alpha)$, then from the proof of Lemma 2.2, α has the form $\alpha = a_1\alpha'$, $\alpha' \in \Sigma^\omega$.

Denote the automaton obtained from $\mathcal{A}$ by removing p_1 and the transition leaving p_1 (labeled a_1) by

$$\mathcal{A}_1 = (Q - \{p_1\}, \Sigma, \delta|_{Q-\{p_1\}}, Q - \{p_1\}, Q - \{p_1\}).$$

If $\mathcal{A}_1$ is not diconnected, then the only source in $\mathcal{A}_1$ is $[p_2] = \{p_2\}$. Suppose that the only transition from p_2 is $p_2 \xrightarrow{a_2} p_3$, $p_3 \in Q - \{p_1, p_2\}$, $a_2 \in \Sigma$ and put

$$\mathcal{A}_2 = (Q - \{p_1, p_2\}, \Sigma, \delta|_{Q-\{p_1, p_2\}}, Q - \{p_1, p_2\}, Q - \{p_1 p_2\}).$$

If $\mathcal{A}_2$ is not diconnected, then we continue our procedure. Obviously, after a finite number of steps, say $k \geq 1$, we get a diconnected $\mathcal{A}_k$. Moreover, $L(\mathcal{A}_k)$ is infinite since R is. It remains to show that $G(\mathcal{A}_k)$ is a cycle. $G(\mathcal{A}_k)$ contains at least one cycle; suppose that there are two distinct cycles (meaning that none of them is contained in the other), the second one being $p_k \xrightarrow{u_k} p_k$. As mentioned, α must start with $a_1a_2\ldots a_{k-1}$. We have that $a_1a_2\ldots a_{k-1}w_k$, $a_1a_2\ldots a_{k-1}u_k \in Fact(\alpha)$. As $a_1a_{2\ldots a_{k-1}}$ appears only at the beginning of α, it follows that w_k is a prefix of u_k or conversely, a contradiction. ■

Because, given a regular language R, a strongly minimal automaton for R is effectively constructable by Lemma 2.1 and it is decidable whether or not an arbitrary finite automaton is diconnected as well as ultimately periodic, we obtain as a consequence of Lemma 2.3 the main result of this section.

THEOREM 2.4: *It is decidable whether or not an arbitrary regular language is in the family* $\mathcal{F}_{fact}$.

3. THE CASE OF BI-INFINITE WORDS

A *bi-infinite* (or two-sided) word is an infinite word without any end. (Usually, an one-sided infinite word is viewed as a function $\alpha : \mathbb{N} \to \Sigma$, Σ

being an alphabet. We can define a bi-infinite word as a function $\alpha : \mathbb{Z} \to \Sigma$ or, in fact, as an equivalence class of the set $\Sigma^{\mathbb{Z}}$ with respect to the equivalence relation defined for $\alpha, \beta \in \Sigma^{\mathbb{Z}}$ by $\alpha \sim \beta$ if and only if there is an integer k such that for any $n \in \mathbb{Z}$, $\alpha(n) = \beta(n+k)$.)

We denote by ${}^{\omega}\Sigma^{\omega}$ the set of all bi-infinite words over Σ. For a finite (non-empty) word $w \in \Sigma^{+}$, we denote by ${}^{\omega}w$ the infinite (to the left) word ${}^{\omega}w = \ldots www$.

Also, we denote by $\mathcal{F}^{bi}_{fact}$ the family of languages of the form $Fact(\alpha)$, for an arbitrary bi-infinite word α, that is,

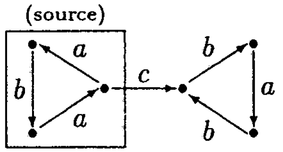
$$\mathcal{F}^{bi}_{fact} = \{L | \text{there are } \Sigma \text{ and } \alpha \in {}^{\omega}\Sigma^{\omega} \text{ with } L = Fact(\alpha)\}.$$

As in the case of one-sided infinite words, it is very easy to prove that it is *undecidable* whether or not an arbitrary context-free language $L \subseteq \Sigma^*$ is in the family $\mathcal{F}^{bi}_{fact}$ or not (the proof uses the undecidability of the problem of whether $L = \Sigma^*$ or not and is similar to the one of Theorem 6 in [MaPa1]).

In what concerns regular languages, we show in this section that the above problem is decidable.

First, notice that things are different from the case of one-sided infinite words; for instance, we can find a regular language $R \in \mathcal{F}^{bi}_{fact}$ such that its automaton $\mathcal{A}$ constructed using Lemma 2.1 has a non-trivial source (*see* the picture below)

$\mathcal{A}$:



Obviously,

$$R = Fact((aba)^* c (bab)^*) = Fact(\alpha) \text{ for } \alpha = {}^{\omega}(aba)\, c\, (bab)^{\omega}.$$

LEMMA 3.1: *For a regular language $R \subseteq \Sigma^*$, $R \in \mathcal{F}^{bi}_{fact}$, let $\mathcal{A} = (Q, \Sigma, \delta, Q, Q)$ be the strongly minimal automaton constructed using Lemma 2.1 for R. Then there is a unique equivalence class $[p] \in Q/_{\equiv}$ which is a source. Moreover, $[p]$ is not trivial.*

Proof: Take $\alpha \in {}^{\omega}\Sigma^{\omega}$ such that $R = Fact(\alpha)$.

Suppose that $[p] \in Q/_{\equiv}$ is a source. Then, we have a transition $p \xrightarrow{a} q$, $a \in \Sigma$, $q \in Q - [p]$, and a word $w = w'aw'' \in R$ such that $\mathcal{A}$ must read $p \xrightarrow{a} q$

in order to accept w. Since α is bi-infinite, w can be prolonged arbitrarily long to the left in R, hence the language accepted by the subautomaton

$$\mathcal{A}^{[p]} = ([p], \Sigma, \delta|_{[p]}, [p], [p])$$

must be infinite. It follows that $[p]$ is not trivial.

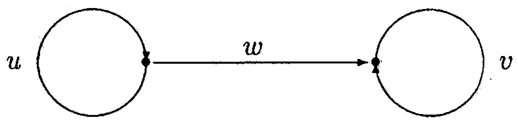
Suppose now that there are two different sources $[p_1], [p_2] \in Q/_{\equiv}$ and the respective transitions and words as above are:

$$p_1 \xrightarrow{a_1} q_1,\ a_1 \in \Sigma,\ q_1 \in Q - ([p_1] \cup [p_2]),\ w_1 = w_1' a_1 w_1'' \in R,$$
$$p_2 \xrightarrow{a_2} q_2,\ a_2 \in \Sigma,\ q_2 \in Q - ([p_1] \cup [p_2]),\ w_2 = w_2' a_2 w_2'' \in R.$$

(So, for $i = 1, 2$, $\mathcal{A}$ must read $p_i \xrightarrow{a_i} q_i$ in order to accept w_i.)

Since w_1 and w_2 appear as factors of α and $[p_1]$ and $[p_2]$ are sources, the occurrences of $w_1' a_1$ and $a_2 w_2''$ are not overlapped and so are the occurrences of $w_2' a_2$ and $a_1 w_1''$. Consequently, the occurrences of w_1 and w_2 in α are not overlapped and we can find, for instance, $w_1 v w_2 \in Fact(\alpha) = R$, $v \in \Sigma^*$. But now w_2 can be accepted by $\mathcal{A}$ without reading $p_2 \xrightarrow{a_2} q_2$, a contradiction. The lemma is proved. ■

LEMMA 3.2: *For a regular language $R \subseteq \Sigma^*$, consider a strongly minimal automation $\mathcal{A} = (Q, \Sigma, \delta, Q, Q)$ accepting R. Then $R \in \mathcal{F}^{bi}_{fact}$ if and only if either $\mathcal{A}$ is diconnected or there are $u, v, w \in \Sigma^*$ such that $G(\mathcal{A})$ has the form*



Proof: If $\mathcal{A}$ is diconnected, then we can prove as in Lemma 2.3 that $R \in \mathcal{F}^{bi}_{fact}$.

If $G(\mathcal{A})$ has the form in the statement, then $R = Fact(\alpha)$ for $\alpha = {}^{\omega}uwv^{\omega} \in {}^{\omega}\Sigma^{\omega}$.

Conversely, suppose that $R \in \mathcal{F}^{bi}_{fact}$. If $\mathcal{A}$ is not diconnected, then, by Lemma 3.1, we get a $p \in Q$ such that $[p]$ is the only source in $\mathcal{A}$ and $[p]$ is not trivial. If

$$\mathcal{A}^{[p]} = ([p], \Sigma, \delta_{[p]}, [p], [p]),$$

then, as in the proof of Lemma 2.3, one can show that

(i) $G(\mathcal{A}^{[p]})$ is a cycle (u in the figure above),

(ii) $G(\mathcal{A}^{Q-[p]})$ is either of the form in Figure 1 or a cycle.

In both cases, the form of $\mathcal{A}$ in the statement of our lemma is obtained. ■

The main theorem of this section is a consequence of Lemma 3.2.

THEOREM 3.3: *It is decidable whether or not an arbitrary regular language is in the family* $\mathcal{F}^{bi}_{fact}$.

Let us further remark that the same result as in Theorem 2.4 holds for left-infinite words as well. Therefore, using also Theorem 3.3, we obtain that it is decidable whether or not an arbitrary regular language is the set of factors of a left-, right-, or bi-infinite word.

REFERENCES

[AuGa] J.-M. AUTEBERT and J. GABARRO, Iterated GSM's and Co-CFL, *Acta Informatica*, 1989, *26*, pp. 749-769.

[AFG] J.-M. AUTEBERT, P. FLAJOLET and J. GABARRO, Prefixes of infinite words and ambiguous context-free languages, *Inf. Process. Lett.*, 1987, *25*, pp. 211-216.

[Bea1] D. BEAUQUIER, Thin homogeneous sets of factors, *Proc. of the 6th Conference of Foundations of Software Technology and Theoret. Comput. Sci.*, Lecture Notes in Computer Science 241, 1986, pp. 239-251.

[Bea2] D. BEAUQUIER, Minimal automaton of a rational cover, *Proc. ICALP'87*, Lecture Notes in Computer Science 267, 1987, pp. 174-189.

[BeaN] D. BEAUQUIER and M. NIVAT, About rational sets of factors of a bi-infinite word, *Proc. ICALP'85*, Lecture Notes in Computer Science 194, 1985, pp. 33-42.

[Ber] J. BERSTEL, Every iterated morphism yields a Co-CFL, *Inf. Process. Lett.*, 1986, *22*, pp. 7-9.

[BoNi] L. BOASSON and M. NIVAT, Adherence of languages, *J. Comput. Syst. Sci.*, 1980, *20*, pp. 285-309.

[Ei] S. EILENBERG, *Automata, Languages and Machines*, Vol. 1, Academic Press, New York, 1974.

[HoUl] J. E. HOPCROFT and J. D. ULLMAN, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley, Reading, Mass., 1979.

[Lo] M. LOTHAIRE, *Combinatorics on Words*, Addison-Wesley, Reading, Mass., 1983.

[MaPa1] S. Marcus and Gh. Păun, Infinite (almost periodic) words, formal languages and dynamical systems, *The Bulletin of the EATCS*, 1994, *54*, pp. 224-231.

[MaPa2] S. Marcus and Gh. Păun, Infinite words and their associated formal language, in: C. Calude, M. J. J. Lennon, H. Maurer (eds.), *Salodays in Auckland* (Auckland Univ. Press, 1994), pp. 95-99.

[Sa] A. Salomaa, *Formal Languages*, Academic Press, New York, 1973.