Journal of Universal Computer Science, vol. 5, no. 9 (1999), 532-541
submitted: 10/6/99, accepted: 15/7/99, appeared: 28/9/99 © Springer Pub. Co.

Decidable and Undecidable Problems of Primitive Words, Regular and Context-Free Languages

Sándor Horváth
Department of Computer Science, L. Eötvös University, Hungary
horvath@cs.elte.hu

Masami Ito
Faculty of Science, Kyoto Sangyo University, Japan
ito@ksuvx0.kyoto-su.ac.jp

Abstract: For any language L over an alphabet X, we define the root set, $root(L)$ and the degree set, $deg(L)$ as follows: (1) $root(L) = \{p \in Q \mid \exists i, i \geq 1, p^i \in L\}$ where Q is the set of all primitive words over X, (2) $deg(L) = \{i \mid \exists p \in Q, p^i \in L\}$. We deal with various decidability problems related to root and degree sets.
Key Words: Root sets, Degree sets, Regular languages, Context-free languages, Decidability problems

1 Introduction

Throughout this paper, X denotes a (finite) alphabet having at least two letters. By X^+ and X^* we denote the free semigroup and the free monoid generated by X, respectively. Morover, λ denotes the empty word over X, i.e., the identity of X^*. By $|u|$ we denote the length of a word $u \in X^*$ ($|\lambda| = 0$). For a grammar G, $L(G)$ denotes the language gennerated by G. Regarding definitions and notatioins concerning formal languages and automata, not defined in this paper, refer, for instance, to [Gi66], [Sa73], [Ha78] and [HoUl79]. A word $u \in X^*$ is said to be *primitive* if $u \neq \lambda$ and $u \neq p^i$ for any $i \geq 2$ and $p \in X^+$; all other elements of X^* are called *nonprimitive.* By $Q(X)$ (or simply by Q if X is fixed) we denote the set of all primitive words over X. The following Theorem A plays an important role in combinatorics of words, and we will use it in the sequel, too.

Theorem A (N.J. Fine and H.S. Wilf, see, e.g., in [Ha78] or [Lo84]) *Let $x, y \in X^+$. If two powers, x^i and y^j $(i, j > 0)$ of x and y, respectively, have a common prefix (or a common suffix) of length $|x| + |y| - gcd(|x|, |y|)$ (here gcd means the greatest common divisor), then for some $z \in X^+$, $x, y \in z^+$.*

It is a known important consequence of Theorem A that any word $u \in X^+$ can uniquely be written in the form $u = p^i$ where $p \in Q$ and $i \geq 1$. Therefore the functions $root : X^+ \to Q(X)$ and $deg : X^+ \to \{1, 2, \ldots\}$ can be defined by putting $root(u) = p$ and $deg(u) = i$, respectively (where p and i are taken from the above, unique representation $u = p^i$ of u). For any $u \in X^+$, $root(u)$ and $deg(u)$ are called the "root of u" and the "degree of u", respectively. So clearly $u \in X^+$ is primitive iff $deg(u) = 1$. We also need the following Theorem B, in the sequel too.

Theorem B (H.-J. Shyr and G. Thierrin, see, e.g., in [Sh91]): *Let $x, y, u, v \in X^+$ such that $x = uv$ and $y = vu$, i.e., x and y are (nontrivial) cyclic permutations – or (nontrivial) conjugates, this term is also used – of one another. Then for any $i > 1$, there exists a $z \in X^+$ such that $x = z^i$ iff there exists a $z' \in X^+$ such that $y = (z')^i$ (z', if exists, is a cyclic permutation of z). Therefore $x \in Q$ iff $y \in Q$.*

By Theorem B, the value of the function *deg* is *invariant under cyclic permutation* of its argument. From this it directly follows that the sets Q, $X^* \setminus Q$ and $X^+ \setminus Q$ are closed under cyclic permutation of words.

The functions *root* and *deg* can be extended without any difficulty - as is done in [Ho95] -, to X^* by putting $root(\lambda) = \lambda$ and $deg(\lambda) = 0$. However, in this paper we need only the "natural extensions" of *root* and *deg*, from words to languages, as follows. For any $L \subseteq X^*$, we define the *root set* and *degree set* of L, as $root(L) = \{root(w) \mid w \in L\} = \{p \in Q \mid \exists i, i \geq 1, p^i \in L\}$ and $deg(L) = \{deg(w) \mid \ w \in L\} = \{i \mid \exists p \in Q, p^i \in L\}$.

We recall that a language is called *bounded* if there are nonempty words $w_1, \ldots, w_k$ such that, $L \subseteq w_1{}^* \ldots w_k{}^*$ (e.g., the languages $L = \emptyset$ and $\{a^n b^n \mid n \geq 1\}$ are bounded languages), see, e.g., in [Gi66]. Concerning bounded languages we will use the following.

Theorem C (Theorem 5.5.2 in [Gi66]) *(a) It is decidable for an arbitrary context-free grammar G, whether $L(G)$ is bounded. (b) If the answer in point (a) is "yes" then (nonempty) words $w_1, ..., w_k$ can be constructed so that, $L(G) \subseteq w_1{}^* ... w_k{}^*$.*

Finally we mention a few related, earlier papers, [DoHoIt91], [DoHoIt93], [ItKaShYu88], [DoHoItKaKa93], [DoHoItKaKa94], [Pe96] and [HoKu95], on primitive words, for the interested reader. We also mention the following, still unsettled conjecture, which was first formulated in [DoHoIt91].

Conjecture *Q is not context-free.*

2 Results concerning root sets

In this section, we provide various results concerning root sets. The following theorem was proved in [HoKu95], using the elaborate machinery of L system theory.

Theorem 2.1 ([HoKu95]) *It is decidable for an arbitrary regular grammar G, whether $L(G) \subseteq Q$.*

However, we think that it is worthwile to give here an elementary proof for this theorem, as follows.

Lemma 2.1 *Let $L \subseteq X^*$ be a regular language. If $L \subseteq (X^+ \setminus Q)$, then $root(L)$ is finite.*

Proof By Theorem 1 in [ItKa91], L can be represented as $L = L_1 \cup L_2$ where L_1 and L_2 are both regular and $L_1 \subseteq Q^{(2)}$ and $L_2 \subseteq \cup_{i \geq 3} Q^{(i)}$ where $Q^{(j)} = \{q^j \mid q \in Q\}$ for $j, j \geq 2$. Moreover, $root(L_2)$ is finite. Therefore, to prove the lemma, it is enough to show that L_1 is finite. Suppose L_1 is infinite. Let $u^2 \in L_1$ where $u \in Q$ and $|u|$ is large enough. Since L_1 is regular, there exists a decomposition of u, i.e., $u = u_1u_2u_3$ such that $u_2, u_3 \in X^+$ and $(u_1u_2^tu_3)(u_1u_2u_3) \in L_1$ for any $t, t \geq 0$. Thus $(u_1u_2^3u_3)(u_1u_2u_3) \in L_1$. Notice that $(u_1u_2^3u_3)(u_1u_2u_3) \in Q^{(2)}$ and hence $(u_3u_1u_2)^2u_2^2 \in Q^{(2)}$. However, by Schützenberger's theorem (see, for instance, [Lo84, Sh91]), $(u_3u_1u_2)^2u_2^2 \in Q$, a contradiction. Hence L_1 must be finite. This comletes the proof of the lemma.

Lemma 2.2 *Let $L \subseteq X^*$ be regular and let L be accepted by a finite deterministic automaton $\mathbf{A} = (S, X, \delta, s_0, F)$ with $|S| = n$ where $|S|$ denotes the cardinality of S. If $root(L)$ is regular and $f^2f^* \cap L \neq \emptyset$ for $f \in Q$, then $|f| \leq n^n$.*

Proof Since $root(L)$ is regular, $L \setminus root(L)$ is regular. By Lemma 2.1, $K := \{g \in Q \mid g^2g^* \cap (L \setminus root(L)) \neq \emptyset\}$ is finite. Let $f \in K$ with $|f| = max\{|g| \mid g \in K\}$. Let $f^m \in L$ where $m \geq 2$. We can assume that $2 \leq m \leq n+1$. Suppose $|f| > n^n$.Then f can be represented as $f = f_1f_2f_3$ where $f_2, f_1f_3 \in X^+$ and $\delta(s_0, f^tf_1) = \delta(s_0, f^tf_1f_2)$ for any $t, 0 \leq t \leq m-1$. This implies that $(f_1f_2^if_3)^m \in L$ for any i, $i \geq 1$. By the maximality of $|f|$, $f_1f_2^if_3 \notin Q$ for any i, $i \geq 1$. Hence $f_2^if_3f_1 \notin Q$ and $f_2^2f_2^if_3f_1 \notin Q$. Let $f_2^if_3f_1 = g^j$ where $g \in Q$ ang $j \geq 2$. Then $f_2^2f_2^if_3f_1 = f_2^2g^j \notin Q$. By Schützenberger's theorem, $f_2 = g^k$ for some k,$k \geq 1$. Since $g^j = f_2^if_3f_1 = f_2^{i-1}f_2f_3f_1 = g^{k(i-1)}(f_2f_3f_1)$ and $f_3f_1 \neq 1$, $f_2f_3f_1 = g^d$ for some $d, d \geq 2$. Hence $f_2f_3f_1 \notin Q$ and $f_1f_2f_3 \notin Q$, a contradiction. Consequently, $|f| \leq n^n$.

Lemma 2.3 *Let $L \subseteq X^*$ be regular and let L be accepted by a finite deterministic automaton $\mathbf{A} = (S, X, \delta, s_0, F)$ with $|S| = n$. If $L \cap (L^+ \setminus Q) \neq \emptyset$, then there exists $f \in Q$ with $|f| \leq n^n$ such that $f^2f^* \cap L \neq \emptyset$.*

Proof By assumption, there exist $g \in Q$ and $i, i \geq 2$ such that $g^i \in L$. Since L is accepted by $\mathbf{A}$, we can assume that $2 \leq i \leq n+1$. If $|g| \leq n^n$, we are done. Now let $|g| > n^n$. Then g can be represented as $g = g_1g_2g_3$ where g_2, $g_1g_3 \in X^+$ and $\delta(s_0, g^tg_1) = \delta(s_0, g^tg_1g_2)$ for any t, $0 \leq t \leq i-1$. This implies that $(g_1g_3)^i \in L$.

Remark that $0 < |g_1 g_3| < |g|$. Hence there exists $g' \in Q$ such that $0 < |g'| < |g|$ and ${g'}^2 g' \cap L \neq \emptyset$. If $|g'| > n^n$, we continue the same procedure. Finally, we can obtain some $f \in Q$ with $|f| \leq n^n$ such that $f^2 f^* \cap L \neq \emptyset$.

Proof of Theorem 2.1 Notice that a finite deterministic automaton $\mathbf{A} = (S, X, \delta, s_0, F)$ that accepts $L(G)$ can be effectively constructed from the regular grammar G. Let $|S| = n$ and let $H = \{h \in Q \mid |h| \leq n^n\}$. By Lemma 2.3, if $(\cup_{h \in H} h^2 h^*) \cap L(G) \neq \emptyset$, then $L(G) \cap (X^+ \setminus Q) \neq \emptyset$. Otherwise, $L(G) \subseteq Q$. This completes the proof of the proposition.

Theorem 2.2 *It is decidable for an arbitrary regular grammar G, whether $root(L(G))$ is regular.*

Proof Let $L(G) \subseteq X^*$ be a regular language that is accepted by a finite deterministic automaton $\mathbf{A} = (S, X, \delta, s_0, F)$ with $|S| = n$. Let $H = \{h \in Q \mid |h| \leq n^n\}$. Consider $L_H = L(G) \setminus (\cup_{h \in H} h^2 h^*)$. Obviously, L_H is regular. By Lemma 2.2, $root(L(G))$ is regular iff $L_H \subseteq Q$. From Theorem 2.1, it follows that it is decidable whether $root(L(G))$ is regular.

Corollary 2.1 *It is decidable for an arbitrary regular grammar G, whether $root(L(G))$ is finite.*

Proof The theorem follows from the fact that it is decidable for a regular language L, whether L is finite.

Theorem 2.2 can be considered as a generalization of Corollary 2.1. Now we prove another generalization of Corollary 2.1.

Theorem 2.3 It is decidable for an arbitrary context-free grammar G, whether $root(L(G))$ is finite.

Proof If $root(L(G))$ is finite and nonempty, say, $root(L(G)) = \{p_1, ..., p_m\} (\subseteq Q), m \geq 1$, then we have $L(G) \subseteq {p_1}^* \cup \cdots \cup {p_m}^* \subseteq {p_1}^* \ldots {p_m}^*$, so in this case $L(G)$ is a bounded language. We recall that a language L is called *bounded* if there are nonempty words $w_1, ..., w_k$ such that, $L \subseteq {w_1}^* \ldots {w_k}^*$ – e.g., the empty language $L = \emptyset$ is trivially a bounded language – (see, e.g., in [Gi66]). Therefore, if $L(G)$ is not bounded then $root(L(G))$ is necessarily infinite. Furthermore, by using Theorem C, we can give the following effective procedure for deciding whether $root(L(G))$ is finite. If in applying point (a) of Theorem C, the answer is "no" then $root(L(G))$ is infinite. If, however, this answer is "yes", i.e., $L(G)$ is bounded, then in point (b) we construct a suitable sequence of (nonempty) words $w_1, ..., w_k$. We can suppose that the sequence $w_1, ..., w_k$ consists of primitive words (otherwise we can replace it by the sequence $root(w_1), ..., root(w_k)$, furthermore, in this new sequence equal neighbouring terms can be identified), and clearly that even the set $\{w_1, \ldots, w_k\}$ is closed under cyclic permutation (by this we mean that if $w \in \{w_1, \ldots, w_k\}$ and w' is a cyclic permutation of w, then also $w' \in \{w_1, \ldots, w_k\}$). Now, if

$$L' := L(G) \setminus ({w_1}^* \cup \cdots \cup {w_k}^*)$$

is finite, then of course, $root(L(G))$ is finite, too. The finiteness of the (context-free) language L' is clearly decidable, since a context-free grammar G' for L' can effectively be constructed from G, etc. If, however, in the former decision procedure the language L' proves to be infinite, then we can show that $root(L')$ is infinite, too, which, of course, implies that also $root(L(G))$ is infinite. Suppose now indirectly that L' is infinite but $root(L)$ is finite, say,

$$root(L') = \{p_1, \ldots, p_s\} \subseteq Q.$$

Then the following properties (1) - (6) must simultaneously hold:

(1) $w_1, \ldots, w_k, p_1, \ldots, p_s \in Q$,

(2) $\{w_1, \ldots, w_k\}$ is closed under cyclic permutation,

(3) $\{w_1, \ldots, w_k\} \cap \{p_1, \ldots, p_s\} = \emptyset$,

(4) $L' = L(G) \setminus ({w_1}^* \cup \cdots \cup {w_k}^*) \subseteq L(G) \subseteq {w_1}^* \ldots {w_k}^*$,

(5) $L' \subseteq {p_1}^* \cup \cdots \cup {p_s}^*$,

(6) L' is infinite.

This is, however, impossible, by Theorems A and B.

Now, let

(I) $m := max\{|p_1|, \ldots, |p_s|\}$.

By (6) there is a $t \in L'$ with

(II) $|t| \geq (|w_1| + m - 1) + \cdots + (|w_k| + m - 1)$.

By (4) , t is of the form $t = {w_1}^{r_1} \ldots {w_k}^{r_k}$. By (II) and the pigeonhole principle, there is a $j \in \{1, \ldots, k\}$ such that,

(III) $|w_j{}^{r_j}| \geq |w_j| + m - 1$

By (5) there are $n \in \{1, \ldots, s\}$ and $e \geq 1$ such that, $t = p_n{}^e$. So $w_j{}^{r_j}$ is a subword of $p_n{}^e$, and by (I) and (III) we have

(IV) $|w_j{}^{r_j}| \geq |w_j| + |p_n| - 1 \geq |w_j| + |p^n| - gcd(|w_j|, |p_n|)$.

Furthermore, by Theorem B,

(V) there is a cyclic permutation p'_n of p_n and an $e' \geq 2$ such that, $p'_n \in Q$ and $w_j{}^{r_j}$ is a *prefix* of ${p'_n}^{e'}$, and so

(VI) $|{p'_n}^{e'}| \geq |w_j{}^{r_j}| \geq |w_j| + |p_n| - gcd(|w_j|, |p_n|) = |w_j| + |p'_n| - gcd(|w_j|, |p'_n|)$.

Now, by (VI) and Theorem A, there should be a nonempty word z such that, $w_j, p'_n \in z^+$, so by (1) , $p'_n = w_j$, and by (V) and (2), even $p_n \in \{w_1, \ldots, w_k\}$, i.e., we should have $p_n \in \{w_1, \ldots, w_k\} \cap \{p_1, \ldots, p_s\}$, in contradiction with (3).

Notice that Theorem 2.3 implies Corollary 2.1 as well. Now we consider the context-freeness of root sets.

Theorem 2.4 *The problem, whether $root(L(G))$ is context-free for an arbitrary context-free grammar G, is undecidable (or not even partially decidable).*

Proof Let $\alpha = \{(u_i, v_i) \mid u_i, v_i \in \{a, b\}^+, i = 1, \ldots, n\}$ $(n \geq 1)$ be an (instance of the) PCP (*Post Correspondence Problem*) on the alphabet $\{a, b\}$. We recall that a solution of α is a finite, nonempty sequence $(i_1, \ldots, i_k) \in \{1, \ldots, n\}^+$ such that, $u_{i_1} \ldots u_{i_k} = v_{i_1} \ldots v_{i_k}$. It is a well-known result that the problem, whether an arbitrary PCP α (on the alphabet $\{a, b\}$) has a solution, is undecidable (see, e.g., in [Sa73]). By a simple, refined analysis of this undecidability result, it can easily be seen that the set of *$PCP's$ α having no solution,* is *not even recursively enumerable.* Now, to an arbitrary PCP α we assign the following three context-free languages (the first two of which we take from the above mentioned book [Sa73]):

$$L_{\alpha,u} := \{u_{i_1} \ldots u_{i_k} c a^{i_k} b \ldots a^{i_1} b \mid k \geq 1; i_1, \ldots, i_k \in \{1, \ldots, n\}\},$$

$$L_{\alpha,v} := \{v_{i_1} \ldots v_{i_k} c a^{i_k} b \ldots a s^{i_1} b \mid k \geq 1; i_1, \ldots, i_k \in \{1, \ldots, n\}\},$$

$$L(\alpha) := L_{\alpha,u} c L_{\alpha,v} c^2 (L_{\alpha,v} c L_{\alpha,v} c^2)^+.$$

It is easy to see, that a context-free grammar for each of the languages $L_{\alpha,u}$, $L_{\alpha,v}$ and $L(\alpha)$ can simply be constructed from α, in an effective way, and that $L_{\alpha,u} \cap L_{\alpha,v} = \emptyset$ iff α has no solution. Therefore, if α has no solution, then $L(\alpha) \subseteq Q$ holds, so in this case $root(L(\alpha)) = L(\alpha)$, context-free. If, however, α has a solution, then $root(L(\alpha))$ is not context-free because intersecting it with the regular language

$$R := (\{a,b\}^+c)^4 c,$$

we get a non-context-free (but context-sensitive) language:

$$root(L(\alpha)) \cap R = \{(u_{i_1} \dots u_{i_k} c a^{i_k} b \dots a^{i_1} bc)^2 c \mid (i_1, \dots, i_k) \text{ is a solution of } \alpha\}.$$

(The non-context-freeness of this language is easily seen by using the Bar-Hillel lemma.) So we have: $root(L(\alpha))$ is context-free iff α has no solution, and this proves the theorem.

Theorem 2.5 *Let $|X| \geq 3$. The problem, whether $root(L(G))$ is regular for an arbitrary context-free grammar G, is undecidable (or not even partially decidable).*

To prove this, we need the following lemma that can be easily shown.

Lemma 2.4 *Let $c \in X$ and let $Y = X \setminus \{c\}$. Let L be a context-free language over Y. Then L is regular if and only if cL is regular.*

Proof of Theorem 2.6 Suppose that the problem in Theorem is decidable. Let c and Y be above-mentioned ones. Let L be any context-free language over Y. Then cL is a context-free language over X. Now consider $root(cL)$. Then $root(cL) = cL$. By assumption, in this case, we can decide whether cL is a regular language over X. By Lemma 2.4, we can decide whether L is a regular language over Y. However, it is known that the latter problem is undecidable. Therefore, the problem in Theorem should be undecidable. Since the all considered undecidable problem can be deducted to the PCP problem, the problems are also not even partially decidable.

Remark 2.1 In fact, making an appropriate coding on 2-letter alphabet, we can show that the above theorem holds true for $|X| = 2$ as well.

3 Results concerning degree sets

In this section, we deal with two decidability problems of degree sets. The following lemma can be proved similarly as Lemmas 2.2 and 2.3.

Lemma 3.1 *If A is a finite deterministic automaton having n (≥ 1) states, $k \geq 1$, X is the alphabet of A, $z^k \in L(A)(\subseteq X^*), |z| > n^n$, then there exist $x, u, y \in X^*$ with $xuy = z, |u| > 0, |xy| > 0$, such that $(xy)^k \in L(A)$.*

Theorem 3.1 *It is decidable for an arbitrary regular grammar G, whether $deg(L(G))$ is finite.*

Proof By definition, $deg(L(G)) = \{k \geq 0 \mid p^k \in L(G)$ for some $p \in Q\}$. If here, in a power $p^k \in L(G)$ we have $k \geq 1$ and $|p| \geq n^n$ where n is the number of states of an automaton A accepting $L(G)$, i.e., $L(A) = L(G)$ – such an A can effectively be constructed from G, then by Lemma 3.1, there is a $w_1 \in X^+$ with $|w_1| < |p|$ such that ${w_1}^k \in L(G)$. Here the word w_1 is not necessarily primitive. If still $|w_1| > n^n$, then by applying Lemma 2.1 again, we obtain a word w_2 with $1 \leq |w_2| < |w_1|$ such that, ${w_2}^k \in L(G)$, and so on, and finally we get a word w_r for which $1 \leq |w_r| \leq n^n$ and ${w_r}^k \in L(G)$. This implies that k $deg(w_r) \in deg(L(G))$ because $w_r = root(w_r)^{deg(w_r)}$. Therefore from the evident estimations $1 \leq deg(w_r) \leq |w_r| \leq n^n$ for all possible w_r, it simply follows that $deg(L(G))$ is finite iff the regular language $L' := L(G) \cap (\cup_{w \in X^+, |w| \leq n^n} w^+)$ is finite, and having G, we can effectively decide whether L' is finite.

Theorem 3.2 *The problem, whether $deg(L)$ is finite for an arbitrary context-free grammar G, is undecidable (or not even partially decidable).*

Proof We use PCP α and the context-free languages $L_{\alpha,u}$ and $L_{\alpha,v}$ from the proof of Theorem 2.5, and the fact that the set $\{\alpha$, a PCP on $\{a, b\}^+ \mid \alpha$ has no solution$\}$ is not recursively enumerable.

Now for an arbitrary α we define the context-free language

$$L(\alpha)' := L_{\alpha,u}c^2(L_{\alpha,v}c^2)^+.$$

Then we clearly have the following chain of equivalencies:

$deg(L(\alpha)')$ is finite

iff $deg(L(\alpha)') = \{1\}$

iff $L(\alpha)' \subseteq Q(\{a, b, c\})$

iff α has no solution.

This implies that the set $\{G$, a context-free grammar $\mid deg(L(G))$ is finite$\}$ is not recursively enumerable.

For the class of regular languages, we can show easily the following result.

Theorem 3.3 *For every regular language L, the set $deg(L)$ is ultimately periodic.*

4 Conclusions

It is clear that all the above studied (or partly, only mentioned) properties of the root and degree sets represent *nontrivial finitely invariant* (shortly *n.f.i.*) properties of the original languages, in both of the classes of *type* 1 and *type* 0 languages, see [DoHoItoKaKa94]. For instance, the properties "$root(L)$ is finite" and "$deg(L)$ is ultimately periodic", as properties of the original language L, are n.f.i. properties of both the *type* 1 and *type* 0 languages. Therefore, by Theorem 5.9 of [DoHoItoKaKa94], if t is such a property, then the quantifier complexity of the decision problem, whether $L(G)$ has property t, is strictly above the $\Sigma_1 - \Pi_1$ level in Kleene's arithmetical hierarchy, if G is an arbitrary *type* 1 or *type* 0 grammar. This means that if $\mathbf{G}_i, i = 0, 1$, is the set of *type* i grammars, then neither the set $\mathbf{H}_{i,t} := \{G \in \mathbf{G}_i \mid L(G) \text{ has property } t\}$, nor its complement $\mathbf{G}_i \setminus \mathbf{H}_{i,t}$ is recursively enumerable.

Aknowledgement

This work was supported by the "Automata and Formal Languages" project of the Hungarian Academy of Science and the Japan Society for the Promotion of Science (No. 15), and grants in-Aid for Scientific Research (No. 10440034), Ministry of Education, Science and Culture of Japan.

References

[DoHoIt91] P. Dömösi, S. Horváth and M. Ito, On the connection between formal languages and primitive words, In: Proc. First Sess. Scientific Communication, Analele Univ. din Oradea, Fasc. Matematica (1991), 5 - 67.

[DoHoIt93] P. Dömösi, S. Horváth and M. Ito, Formal languages and primitive words, Publicatione Mathematicae 42 (1993), 315 - 321.

[DoHoItKaKa93] P. Dömösi,S. Horváth, M. Ito, M., L. Kászonyi and M. Katsura, Formal languages consisting of primitive words, Lecture Notes in Computer Science 710 (1993), Springer, 194 - 203.

[DoHoItKaKa94] P. Dömösi, S. Horváth, M. Ito, L. Kászonyi, M. Katsura, Some combinatorial properties of words, and the Chomsky-hierarchy, Words, Languages and Combinatorics II (1994), World Scientific, 105 - 123.

[Gi66] S. Ginsburg, The Mathematical Theory of Context-Free Languages, McGraw-Hill Book Comp., New York, London, etc., 1966.

[Ha78] M.A. Harrison, Introduction to Formal Language Theory, Addison-Wesley, Reading, Mass., 1978.

[Ho95] S. Horváth, Strong interchangeability and nonlinearity of primitive words, In: Proc. Workshop AMiLP'95 (Algebraic Methods in Language Processing) (1995), 173 - 178.

[HoKu95] S. Horváth, M. Kudlek, On the classification and decidability problems of primitive words, Pure Mathematics and Applications 6(1995), 171 - 189.

[HoUl79] J.E. Hopcroft, J.D. Ullman, Introduction to Automata Theory, Languages, and Computation, Addison-Wesley, Reading, Mass., 1979.

[ItKa91] M. Ito, M. Katsura, Context-free languages consisting of non-primitive words, International Journal of Computer Mathematics 40 (1991), 157 - 168.

[ItKaShYu88] M. Ito, M. Katsura, H.J. Shyr, S.S. Yu, Automata accepting primitive words, Semigroup Forum 37 (1988), 45 - 52.

[Lo84] M. Lothaire, Combinatorics on Words, 1st ed.: Addison-Wesley, Reading, Mass., 1983, 2nd ed., Cambridge Univ. Press, 1984.

[Pe96] H. Peterson, On the language of primitivewords, Theoretical Computer Science 161 (1996), 141 - 156.

[Sa73] A. Salomaa, Formal Languages, Academic Press, New York, London, 1973.

[Sh91] H.J. Shyr, Free Monoids and Languages, Lecture Notes, 2nd ed., Hon Min Book Company, Taichung, Taiwan, 1991.