Numeration systems on a regular language

Pierre B. A. Lecomte and Michel Rigo
Institut de Mathématiques, Université de Liège,
Grande Traverse 12 (B 37), B-4000 Liège, Belgium.
plecomte@ulg.ac.be, M.Rigo@ulg.ac.be

September 23, 2002

Abstract

Generalizations of positional number systems in which $\mathbb{N}$ is recognizable by finite automata are obtained by describing an arbitrary infinite regular language according to the lexicographic ordering. For these systems of numeration, we show that ultimately periodic sets are recognizable. We also study translation and multiplication by constants as well as the order-dependence of the recognizability.

1 Introduction

A series of recent papers is devoted to numeration systems [3, 4, 9, 10, 13, 16, 23] and is mainly concerned with the study of the so-called recognizable sets of integers. Roughly speaking, a set of integers is recognizable if the representations of its elements have a very simple syntax, i.e., if they form a regular language.

A common way of representing integers, leading to the so-called positional numeration systems, is to consider a strictly increasing sequence $(U_n)_{n\in\mathbb{N}}$ of positive integers and to use some algorithm (such as the greedy algorithm) to represent each natural number x by a word $c_0\cdots c_n$ such that $c_0U_n+\cdots+c_nU_0=x$ [12]. For example, with $U_n=p^n$ and the greedy algorithm, one gets the standard numeration system with base p.

Among the possibly recognizable sets of integers, $\mathbb{N}$ is of special interest. For instance, if $\mathbb{N}$ is recognizable, then one can easily check whether a word over the alphabet of the digits represents an integer or not. Under quite general assumptions, it is shown in [23] that for $\mathbb{N}$ to be recognizable, it is necessary that $(U_n)_{n\in\mathbb{N}}$ satisfies a linear recurrence relation. In [16], a sufficient condition is given in terms of the polynomials of the recurrence that $(U_n)_{n\in\mathbb{N}}$ satisfies (it generalizes the case studied in [18]). Examples of such systems are the numeration systems defined by a recurrence relation whose characteristic polynomial is the minimum polynomial of a Pisot number (such as the standard numeration systems or the Fibonacci system [4, 14]; a Pisot number is an algebraic integer greater than one, the Galois conjugates of which have modulus less than one).

A nice description of the recognizable sets has been obtained for the systems related to a Pisot number [4, 5]. They are the sets of integers that can be defined in Presburger arithmetic extended by some predicate associated with the considered Pisot number. In particular, various operations do preserve the recognizability, such as addition, translation and multiplication by a constant.

In [6], Cobham shows that the only sets that are simultaneously recognizable with respect to two standard numeration systems having multiplicatively independent integer bases are precisely the finite unions of arithmetic progressions. This remarkable result has been extended to non-standard linear systems [5, 9, 10, 15, 22].

In the above mentioned results, a property of the considered systems seems to play a crucial role: the function r that maps a natural number onto its representation is strictly increasing if one endows the set of words, $r(\mathbb{N})$, with the lexicographic order (this is an assumption in [23]; it is a consequence of the greedy algorithm). Observe that a numeration system having this property is completely determined by the language $r(\mathbb{N})$ of representations of all non-negative integers and the ordering of the digits, the sequence U_n and the algorithm defining the system being just extra data devised to compute the function r in some "practical" fashion.

Taking this into account, we thus define an *(abstract) numeration system* as being a triple $S = (L, \Sigma, <)$ where L is an infinite language over the totally ordered alphabet $(\Sigma, <)$. Enumerating the elements of L lexicographically with respect to $<$ leads to a one-to-one map r_S from $\mathbb{N}$ onto L. To any natural number n, it assigns the $(n+1)^{th}$ word of L, its *S-representation*, while the inverse map val_S sends any word belonging to L onto its *numerical value*. A subset $X \subset \mathbb{N}$ is said to be *S-recognizable* if $\mathrm{r}_S(X)$ is a regular subset of L.

Having in mind a possible generalization of Cobham's theorem, it is natural to check whether the ultimately periodic subsets of $\mathbb{N}$ are S-recognizable. Of course, if they are, then $L = \mathrm{r}_S(\mathbb{N})$ is regular. It is a quite remarkable fact that conversely, if L is regular, then every arithmetic progression is indeed S-recognizable (a special case of this result has been obtained separately in [19]); this is proved below in Theorem 4.

As recalled above, the recognizability of $\mathbb{N}$ is an important property that is often required. Unless otherwise stated, we assume in the sequel that L is a regular language. Under this assumption, we obtain algorithms to compute r_S and val_S. The first is a generalization of the greedy algorithm involving the complexity functions of the languages accepted from the different states of the minimal automaton of L (the complexity function of a language $L \subset \Sigma^*$ maps $n \in \mathbb{N}$ onto $\#(L \cap \Sigma^n)$; for more about minimal automata, see for instance [11]). Both proved to be quite useful in many concrete experiments.

In a positional numeration system, each digit has its own weight so that the question of changing the order of the digits is somewhat irrelevant in this case. In an abstract numeration system, however, the letters have no *a priori* individual weight and, as we show with the help of the language $\{a,b\}^* \setminus a^*b^*$, the family of recognizable sets depends on the ordering of the alphabet. However, we exhibit two classes of regular languages for which the recognizability of a set of integers is independent of the order on the alphabet. One of these classes is the set of slender languages [1]. The other is the set of the languages $L \subset \Sigma^*$ for which the complexity functions related to the languages recognized from the different states of the minimal automaton of L differ only at finitely many places.

As for the stability of the recognizability under natural arithmetic operations, we show that for each t, a subset X of $\mathbb{N}$ is S-recognizable if and only if $X+t$ is S-recognizable. On the other hand, multiplication by a constant generally does not preserve recognizability so that addition is not a regular map as well. For example, in the numeration system S based on the language a^*b^*, the set of $t \in \mathbb{N}$ for which tX is S-recognizable if X is S-recognizable consists of the odd perfect squares. Note that in this case, the function val_S is nothing else but the well-known function of Peano [25]; surprisingly, the proof of the result is difficult and it relies partly on the properties of Pell's equation [8, 24].

2 Basic definitions and notation

In this paper, if Σ is a finite alphabet then Σ^* is the free monoid (with identity ε) generated by Σ. For a set S, $\#S$ denotes the cardinality of S and for a string $w \in \Sigma^*$, $|w|$ denotes the length of w.

Let $L \subset \Sigma^*$ be a regular language. We define $M_L = (K, s, F, \delta, \Sigma)$ to be the minimal automaton of L, where K is the set of states, s is the initial state, F is the set of final states and $\delta : K \times \Sigma \to K$ is the transition function. We often write $k.\sigma$ instead of $\delta(k, \sigma)$.

Recall that the elements of K are the sets $w^{-1}.L = \{v \in \Sigma^* : wv \in L\}$, $w \in \Sigma^*$ (see for instance [11, III.5]). The state k is of the form $w^{-1}.L$ if and only if $k = s.w$. In this case, $w^{-1}.L$ is the set L_k of words accepted by M_L from k, i.e.,

$$L_k = \{x \in \Sigma^* : \delta(k, x) \in F\}.$$

In particular, $L = L_s$.

We denote by $u_l(k)$ the number $\#(L_k \cap \Sigma^l)$ of words of length l belonging to L_k and by $v_l(k)$ the number of words of length at most l belonging to L_k,

$$v_l(k) = \sum_{i=0}^{l} u_i(k).$$

If we are only interested in the number of words belonging to L, then we simply write u_l and v_l instead of $u_l(s)$ and $v_l(s)$, provided that it does not lead to any confusion.

The *lexicographic ordering* is usually understood as being defined for words of the same length. We extend it as follows. Let u and v be two words. We say that $u < v$ if $|u| < |v|$ or if $|u| = |v|$ and there exist letters $\sigma < \gamma$ such that $u = w\sigma u'$ and $v = w\gamma v'$. This ordering is sometimes called "radix order" or "genealogic order".

Definition 1 A *numeration system* S is a triple $(L, \Sigma, <)$ where L is an infinite regular language over the totally ordered alphabet $(\Sigma, <)$.

For each $n \in \mathbb{N}$, $\mathrm{r}_S(n)$ denotes the $(n+1)^{th}$ word of L with respect to the lexicographic ordering and is called the *S-representation* of n.

Observe that the map $\mathrm{r}_S : \mathbb{N} \to L$ is an increasing bijection. For $w \in L$, we set $\mathrm{val}_S(w) = \mathrm{r}_S^{-1}(w)$. We call $\mathrm{val}_S(w)$ the *numerical value* of w.

Definition 2 Let S be a numeration system. A subset X of $\mathbb{N}$ is *S-recognizable* if $\mathrm{r}_S(X)$ is recognizable by a finite automaton.

It is well-known that for the system in base 2 or the Fibonacci system, the set of representations of all non-negative integers is a regular language. Therefore, these systems can also be obtained with our definition of numeration system.

Example 1 Let $\Sigma = \{a, b\}$, $a < b$ and $L = b\{a, b\}^* \cup \{\varepsilon\}$. We use the alphabet $\{a, b\}$ instead of $\{0, 1\}$ because generally we cannot give any weight to the letters (see for instance, Example 3 below). To obtain usual representations, consider the isomorphism h defined by $h(a) = 0$ and $h(b) = 1$. Then $h(L)$ is the set of all representations in base 2 if we do not allow leading 0's. Notice that allowing leading 0's gives another system of numeration; see Table 1. One can see that if we allow leading 0's, the words 0, 00 and 000 have different numerical values.

$\mathbb{N}$	$h(b\{a,b\}^* \cup \{\varepsilon\})$	$h(\{a,b\}^*)$
0	ε	ε
1	1	0
2	10	1
3	11	00
4	100	01
5	101	10
6	110	11
7	111	000
$\vdots$	$\vdots$	$\vdots$

Table 1: Allowing leading 0's changes the system of numeration.

Example 2 As easily seen, in the positional numeration system associated with the Fibonacci sequence $F_{n+2} = F_{n+1} + F_n$ with $F_0 = 1$ and $F_1 = 2$, the representations obtained by the greedy algorithm are exactly the words over $\{0, 1\}$ which do not contain two consecutive 1's. So the Fibonacci system can also be defined by the language $L = \{\varepsilon\} \cup b\{a, ab\}^*$ where the ordered alphabet is $(\{a, b\}, a < b)$. For instance, the numerical value of $baab$ is 6 because 1001 represents $F_3 + F_0$ in the positional numeration system related to $(F_n)_{n\in\mathbb{N}}$ but also because $baab$ is the 7^{th} word of L.

Let $S = (L, \Sigma, <)$ be a numeration system. Each $k \in K$ for which L_k is infinite leads to the numeration system $S_k = (L_k, \Sigma, <)$. The functions r_{S_k} and val_{S_k} are simply denoted by r_k and val_k if the context is clear. If L_k is finite, the functions r_k and val_k are defined as in the infinite case but the domain of the former is restricted to $\{0, \ldots, \#L_k - 1\}$.

Here is another example of numeration system which will be used below.

Example 3 Let $\Sigma = \{a, b\}$, $a < b$ and $L = a^*b^*$. According to the lexicographic ordering, the first words of L are ε, a, b, aa, ab, bb, aaa,.... So, for instance, $\mathrm{r}_S(5) = bb$ and $\mathrm{val}_S(aaa) = 6$. The minimal automaton of L is given in Figure 1. We have $L_s = L = a^*b^*$, $L_t = b^*$ and $L_p = \emptyset$. The sequences associated with the states s and t are $u_n(s) = n+1$, $v_n(s) = \frac{1}{2}(n+1)(n+2)$, $u_n(t) = 1$ and $v_n(t) = n + 1$.

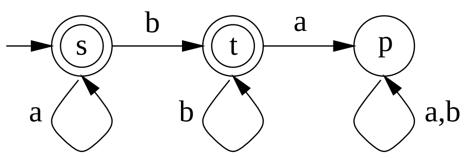

Figure 1: The minimal automaton of a^*b^*.

3 Computation of val_S and recognizability of ultimately periodic sets

In this section, given any numeration system $S = (L, \Sigma, <)$, we indicate how to compute the function val_S and show that the arithmetic progressions $p + \mathbb{N}q$ are S-recognizable.

We first need a lemma.

Lemma 3 *Let $S = (L, \Sigma, <)$ be a numeration system. If $\alpha\beta$ belongs to L_k, $\alpha, \beta \in \Sigma^+$, then*

$$\mathrm{val}_k(\alpha\beta) = \mathrm{val}_{k.\alpha}(\beta) + v_{|\alpha\beta|-1}(k) - v_{|\beta|-1}(k.\alpha) + \sum_{\substack{\alpha' < \alpha \\ |\alpha'| = |\alpha|}} u_{|\beta|}(k.\alpha').$$

Proof. We have to compute the number of words belonging to L_k and lexicographically strictly less than $\alpha\beta$. There are three kinds of such words. The first consists of words of length strictly less than $\alpha\beta$ and counts $v_{|\alpha\beta|-1}(k)$ elements. The second consists of words of length $|\alpha\beta|$ admitting the prefix α. Since a word $\alpha'\beta'$ belongs to L_k if and only if β' belongs to $L_{k.\alpha'}$, we see that there are $\mathrm{val}_{k.\alpha}(\beta) - v_{|\beta|-1}(k.\alpha)$ such words. It is clear that there are

$$\#\{w \in L_k \,:\, w = \alpha'\beta', |\alpha'| = |\alpha|, |\beta'| = |\beta| \text{ and } \alpha' < \alpha\} = \sum_{\substack{\alpha' < \alpha \\ |\alpha'| = |\alpha|}} u_{|\beta|}(k.\alpha')$$

words of the third kind. □

Remark 1 Taking α to be a letter in Lemma 3 leads easily to an effective algorithm to compute val_S.

Remark 2 It follows also from Lemma 3 that for each word $w \in L$, there exist bounded coefficients $c_{k,l} \leq \#\Sigma$ with $k \in K$ and $0 \leq l < |w|$ such that

$$\mathrm{val}_S(w) = \sum_{\substack{k \in K \\ 0 \leq l < |w|}} c_{k,l}\, u_l(k).$$

Theorem 4 *Let $S = (L, \Sigma, <)$ be a numeration system and p, q two non-negative integers. The arithmetic progression $p + \mathbb{N}q$ is S-recognizable.*

The proof below exploits directly the regularity of L. Another proof using quite developed tools (transducers and non-trivial results about numeration systems in an integer base) is given in the appendix at the end of this paper.

Proof. We can assume that $p < q$. We show that the minimal automaton of $\mathcal{A} = \mathrm{r}_S(p + \mathbb{N}q)$ is finite. Its states are the sets

$$w^{-1}.\mathcal{A} = \{x \in \Sigma^* \,:\, \mathrm{val}_S(wx) \equiv p \bmod q\}, w \in \Sigma^*.$$

It is well-known that since L is a regular language, the sequence $(u_n(s))_{n\in\mathbb{N}} = (\#(L \cap \Sigma^n))_{n\in\mathbb{N}}$ satisfies a linear recurrent relation with coefficients in $\mathbb{Z}$ (recall that s is the initial state of M_L). A proof of this result can be found in [2]; it is due to the fact that the series $f_L(X) = \sum_{n\geq 0} u_n(s)\, X^n$ is $\mathbb{N}$-rational. Moreover, $v_{n+1}(s) - v_n(s) = u_{n+1}(s)$. So the sequence $(v_n(s))_{n\in\mathbb{N}}$ is also the solution of a linear recurrence equation and is therefore ultimately periodic in the finite ring $\mathbb{Z}/(q)$, say of period t. By Lemma 3,

$$\mathrm{val}_S(wx) = \mathrm{val}_{s.w}(x) + v_{|wx|-1}(s) - v_{|x|-1}(s.w) + \sum_{\substack{w'<w\\|w|=|w'|}} u_{|x|}(s.w').$$

The language L is regular, so $s.w$ can only take a finite number of values in K (the set of states of M_L). In $\mathbb{Z}/(q)$ and for $|w|$ large enough, the term $v_{|wx|-1}(s)$ can be written as $v_{|x|+i}(s)$ for some $i \in \{0, \ldots, t-1\}$ since $(v_n(s))_{n\in\mathbb{N}}$ is ultimately periodic. Still working in $\mathbb{Z}/(q)$, the term $\sum_{\substack{w'<w\\|w|=|w'|}} u_{|x|}(s.w')$ can be written as $\sum_{k'\in K} j_{k'}\, u_{|x|}(k')$ for some $j_{k'}$'s belonging to $\mathbb{Z}/(q)$.

So, for $|w|$ large enough, any set $w^{-1}.\mathcal{A}$ is of the form

$$\{x \,:\, \mathrm{val}_k(x) + v_{|x|+i}(s) - v_{|x|-1}(k) + \sum_{k'\in K} j_{k'}\, u_{|x|}(k') \equiv p \bmod q\}$$

for some $k \in K$, $j_{k'} \in \{0, \ldots, q-1\}$ and $i \in \{0, \ldots, t-1\}$. So, there are finitely many sets of this kind and the set of states of the minimal automaton of $\mathcal{A}$, $\{w^{-1}.\mathcal{A} : w \in \Sigma^*\}$, is finite. □

4 Computation of r_S and reordering of the alphabet

We now explain how to effectively compute r_S and discuss to what extent the S-recognizable subsets of $\mathbb{N}$ depend on the ordering of the alphabet.

Let $S = (L, \Sigma, <)$ be a numeration system, where $\Sigma = \{\sigma_1 < \cdots < \sigma_p\}$. It is clear that

$$|\mathrm{r}_S(n)| = \inf_m \{m \mid n < v_m\}.$$

Set $|\mathrm{r}_S(n)| = l$; then $n - v_{l-1}$ is the number of words of length l belonging to L and strictly less than $\mathrm{r}_S(n)$.

To determine the first letter of the representation, we have to compute the number N_t^l of words of length l belonging to L and beginning with σ_1 or σ_2 or $\cdots$ or σ_t $(t \le p)$

$$N_t^l = \sum_{i=1}^{t} u_{l-1}(\sigma_i^{-1}.L).$$

If $N_{t-1}^l \le n - v_{l-1} < N_t^l$ then the first letter is σ_t. We proceed in the same way to determine the other letters of the representation. Recall that if k is a state of M_L then $\delta(k, \sigma_j) = \sigma_j^{-1}.k$ (see [11, Theorem 5.2, p. 46]). Hence, the following algorithm computes the S-representation w of a given integer n.

Algorithm 1 *Let l be such that $v_{l-1} \le n < v_l$,*
$k \leftarrow s$
$m \leftarrow n - v_{l-1}$
$w \leftarrow \varepsilon$
for i ranging from 1 *to l do*
    $j \leftarrow 1$
    while $m \ge u_{l-i}[\delta(k, \sigma_j)]$ *do*
        $m \leftarrow m - u_{l-i}[\delta(k, \sigma_j)]$
        $j \leftarrow j + 1$
    $k \leftarrow \delta(k, \sigma_j)$
    $w \leftarrow w\sigma_j$.

As an easy application of Algorithm 1, we obtain our first class of numeration systems in which the recognizable sets are independent of the order of the alphabet.

It is convenient to introduce notation for the change of numeration systems. Given systems $S = (L, \Sigma, <)$ and $T = (L', \Sigma', \prec)$, we set

$$\Theta_{S,T} = \mathrm{r}_T \circ \mathrm{val}_S : L \to L' \text{ and } \Theta'_{S,T} = \mathrm{val}_T \circ \mathrm{r}_S : \mathbb{N} \to \mathbb{N}.$$

If the underlying S and T are known from the context, we simply write Θ and Θ'.

Proposition 5 *Let $S = (L, \Sigma, <)$ and $T = (L, \Sigma, \prec)$ be two numeration systems. Let n_0 be a non-negative integer. If for all states k and k' of $M_L = (K, s, F, \delta, \Sigma)$ and all $n \geq n_0$*

$$u_n(k) = u_n(k'),$$

then $X \subset \mathbb{N}$ is S-recognizable if and only if X is T-recognizable.

Proof. Assume that $\Sigma = \{\sigma_1 < \cdots < \sigma_p\} = \{\sigma_{\nu_1} \prec \cdots \prec \sigma_{\nu_p}\}$ where ν is a permutation of $\{1, \ldots, p\}$. We prove that the graph $\hat{\Theta} = \{(x, y) \in L \times L : \mathrm{val}_S(x) = \mathrm{val}_T(y)\}$ of Θ is regular over the alphabet $\Sigma \times \Sigma$, showing that $\mathrm{r}_T(X) = p_2(\hat{\Theta} \cap p_1^{-1}(\mathrm{r}_S(X))$ is regular if and only if $\mathrm{r}_S(X)$ is regular, where $p_1, p_2 : (\Sigma \times \Sigma)^* \to \Sigma^*$ are the canonical homomorphisms of projection.

Let (x, y) belong to $\hat{\Theta}$. The two systems S and T have the same sequence $(v_n)_{n \in \mathbb{N}}$, thus $|x| = |y|$.

By Algorithm 1, if $|x| \geq n_0$ then

$$x = \underbrace{\sigma_{i_1} \ldots \sigma_{i_l}}_{\alpha} \beta \text{ and } y = \underbrace{\sigma_{i_{\nu_1}} \ldots \sigma_{i_{\nu_l}}}_{\alpha'} \beta'$$

where $|\beta| = |\beta'| = n_0$, $\beta \in L_{s.\alpha}$, $\beta' \in L_{s.\alpha'}$ and

$$\mathrm{val}_{S_{s.\alpha}}(\beta) = \mathrm{val}_{T_{s.\alpha'}}(\beta').$$

To conclude, it is then sufficient to observe that the words of $\hat{\Theta}$ of length at least n_0 are exactly the words accepted by the following nondeterministic finite automaton. The set of states is $(K \times K) \cup \{f\}$. The initial state is (s, s). The new symbol f denotes the unique final state. According to what precedes, there are two kinds of transitions. First, those of label $(\sigma_i, \sigma_{\nu_i})$ mapping the state (k, k') onto $(k.\sigma_i, k'.\sigma_{\nu_i})$. Second, those of label (β, β') mapping (k, k') onto f, provided that $|\beta| = |\beta'| = n_0$, $\beta \in L_k$, $\beta' \in L_{k'}$ and $\mathrm{val}_{S_k}(\beta) = \mathrm{val}_{T_{k'}}(\beta')$. □

Example 4 The language over the alphabet $\{a, b\}$ consisting of the words containing an even number of a's satisfies the hypothesis of Proposition 5.

In the next proposition, we give equivalent formulations of the assumption of Proposition 5. They are expressed in terms of the incidence matrix A_L of the minimal automaton M_L of L. Recall that it is the matrix defined by

$$(A_L)_{i,j} = \sum_{t=1}^{p} \delta_{k_i.\sigma_t, k_j}, \ 1 \leq i, j \leq \kappa,$$

where the σ_t's and the k_i's denote the p letters and the κ states of M_L respectively.

We denote by f_L the characteristic vector of the set of final states:

$$(f_L)_i = \begin{cases} 1, & \text{if } k_i \in F; \\ 0, & \text{otherwise.} \end{cases}$$

Observe that

$$(A_L^m f)_i = \sum_{j=1}^{\kappa} (A_L^m)_{i,j} f_j = u_m(i). \tag{1}$$

Proposition 6 *Let L be a regular language over an alphabet Σ and $M_L = (K, s, F, \delta, \Sigma)$ be its minimal automaton. Let m be the multiplicity of 0 as root of the minimal polynomial of A_L. Let $r > m$. The following assertions are equivalent*

1. $\forall n \geq r$, $\forall k, k' \in K$, $u_n(k) = u_n(k')$,
2. $\forall n \geq m$, $\forall k, k' \in K$, $u_n(k) = u_n(k')$,
3. $\exists \lambda \in \mathbb{N} \setminus \{0\}$*:* $A_L^m f = \lambda\, v$, *with* $v = (1, \ldots, 1)^T$.

In particular, $\forall k \in K$, $\forall i \geq 0$, $u_{m+i}(k) = (\#\Sigma)^i u_m(k)$.

Proof. This follows immediately from (1) and the well-known fact that any polynomial that is annihilated by A_L is the characteristic polynomial of a linear recurrence equation satisfied by each of the sequences $u_n(k)$. □

Here is another easy characterization of the languages for which the assumption of Proposition 5 holds true.

Proposition 7 *Let L be a regular language over an alphabet Σ. Then L satisfies the hypothesis of Proposition 5 if and only if there exist $n_0, \alpha_0 \in \mathbb{N}$ such that for all $w \in \Sigma^*$, $\#((w^{-1}.L) \cap \Sigma^{n_0}) = \alpha_0$.* □

The set of slender languages is the second class of languages for which the recognizable sets of integers are independent of the ordering of the alphabet.

Definition 8 [1] Let d be a positive integer. The language L is said to be *d-slender* if

$$\forall n \geq 0,\ u_n(s) \leq d,$$

L is said to be *slender* if there exists d such that L is d-slender. A regular language is slender iff for some $k \geq 1$ and words x_i, y_i, z_i, $1 \leq i \leq k$,

$$L = \bigcup_{i=1}^{k} x_i\, y_i^* z_i.$$

In this case, L is said to be a *union of single loops* (see [23, 20]).

Lemma 9 [23] *Let L be a regular language over the totally ordered alphabet $(\Sigma,<)$. The set $\mathcal{I}(L,<)$ (resp. $\mathcal{G}(L,<)$) obtained by taking from all the words of L of the same length only the first (resp. last) one in the lexicographic order is regular.* $\square$

Proposition 10 *Let d be a positive integer. Let L be a regular d-slender language. Let $S=(L,\Sigma,<)$ and $T=(L,\Sigma,\prec)$ be two numeration systems. If $X\subset\mathbb{N}$ is S-recognizable then X is T-recognizable.*

Proof. As in the proof of Proposition 5, we show that the graph $\hat{\Theta}$ of the change of systems is regular. Using Lemma 9, we define iteratively the regular languages $I_{i,<}$ and $I_{i,\prec}$ by

$$\left\{\begin{array}{rcl} I_{1,<} & = & \mathcal{I}(L,<)\\ I_{1,\prec} & = & \mathcal{I}(L,\prec), \end{array}\right.$$

and, for $i=2,\ldots,d$,

$$\left\{\begin{array}{rcl} I_{i,<} & = & \mathcal{I}[L\setminus(\bigcup_{j=1}^{i-1} I_{j,<}),<]\\ I_{i,\prec} & = & \mathcal{I}[L\setminus(\bigcup_{j=1}^{i-1} I_{j,\prec}),\prec]. \end{array}\right.$$

Since for all $x\in L$, $|x|=|\Theta(x)|$, the graph of Θ is thus given by

$$\hat{\Theta}=\bigcup_{j=1}^{d}\left[(I_{j,<}\times I_{j,\prec})\cap(\Sigma\times\Sigma)^*\right].\ \square$$

In spite of the two previous propositions, the change of ordering of the alphabet generally does not preserve recognizability as we shall see in the case where $\Sigma=\{a,b\}$ and $L=\Sigma^*\setminus a^*b^*$.

Lemma 11 *Let $n\in\mathbb{N}$. For $U=(\Sigma^*,\Sigma,a<b)$ and $V=(\Sigma^*,\Sigma,b\prec a)$ one has*

$$\Theta'_{U,V}(n)=3.2^l-n-3,$$

where $l=|\mathrm{r}_U(n)|$.

Proof. Observe that since $\#\Sigma^l=2^l$, if $w_1<\cdots<w_{2^l}$ then $w_{2^l}\prec\cdots\prec w_1$. Moreover $2^l-1\le n\le 2^{l+1}-2$. Thus

$$\Theta'(n)=2^{l+1}-2-[n-(2^l-1)].\ \square$$

Proposition 12 *Let $\Sigma = \{a,b\}$ and $L = \Sigma^* \setminus a^*b^*$. For all $n \geq 2$, if $l = |\mathrm{r}_U(n-1)|$ then*

$$\Theta_{S,T}(ba\,b^n) = ab\,a^{n-l-1}\,b\,\mathrm{r}_U(n-1),$$

where $S = (L, \Sigma, a < b)$, $T = (L, \Sigma, b \prec a)$ and $U = (\Sigma^*, \Sigma, a < b)$. In particular, $\mathrm{val}_S(ba\,b^2b^*)$ is not T-recognizable.

Proof. The minimal automaton M_L of L is given on Figure 2. Therefore

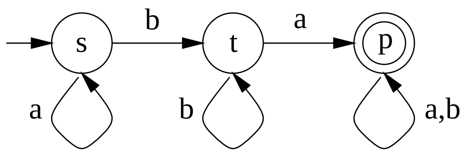

Figure 2: The minimal automaton of $\Sigma^* \setminus a^*b^*$.

$L_p = \Sigma^*$,

$$\begin{cases} u_0(s) = u_1(s) = 0, \\ u_n(s) = 2^n - n - 1, \quad \forall n \geq 2, \end{cases}$$

while $u_n(t) = 2^n - 1$ for all $n \in \mathbb{N}$.

In L, there are $v_{n+1}(s)$ words of length at most $n+1$, $u_{n+1}(s)$ words of length $n+2$ beginning with a and $u_n(p) - 1$ words of length $n+2$ beginning with ba. Hence, the number of words belonging to L and lexicographically less than $ba\,b^n$ is

$$\mathrm{val}_S(ba\,b^n) = \sum_{i=2}^{n+1}(2^i - i - 1) + 2^{n+1} + 2^n - n - 3.$$

Using Lemma 3, we sketch the computation of $\mathrm{val}_T[ab\,a^{n-l-1}\,b\,\mathrm{r}_U(n-1)]$

$$\begin{aligned}
&= \mathrm{val}_t[a^{n-l-1}\,b\,\mathrm{r}_U(n-1)] + \sum_{i=2}^{n+1}(2^i - i - 1) + 2^n + n \\
&= \mathrm{val}_p[a^{n-l-2}\,b\,\mathrm{r}_U(n-1)] + \sum_{i=2}^{n+1}(2^i - i - 1) + 2^{n+1} - 1 \\
&\vdots \\
&= \mathrm{val}_p[b\,\mathrm{r}_U(n-1)] + \sum_{i=2}^{n+1}(2^i - i - 1) + 2^{n+1} - 1 + \sum_{i=l+2}^{n-1} 2^i \\
&= \mathrm{val}_p[\mathrm{r}_U(n-1)] + \sum_{i=2}^{n+1}(2^i - i - 1) + 2^{n+1} - 1 + \sum_{i=l+2}^{n-1} 2^i + 2^l \\
&= \Theta'(n-1) + \sum_{i=2}^{n+1}(2^i - i - 1) + 2^{n+1} - 1 + 2^n - 3.2^l.
\end{aligned}$$

Hence the value of $\Theta_{S,T}(ba\,b^n)$, in view of Lemma 11. Applying the pumping lemma, it is now straightforward to check that $\mathrm{val}_S(ba\,b^2b^*)$ is not T-recognizable. $\square$

5 Translation by a constant

Here we show that the S-recognizability of a set is conserved under translation by a constant. First we recall some results about numeration systems.

Definition 13 If $x = x_n \cdots x_0$ is a word over an alphabet $B \subset \mathbb{Z}$ and $p \in \mathbb{N} \setminus \{0,1\}$. Then we define $\pi_p(x)$ as

$$\pi_p(x) = \sum_{i=0}^{n} x_i\, p^i.$$

Lemma 14 [13] *Let $p \in \mathbb{N} \setminus \{0,1\}$. The normalization function*

$$\nu : \{1, \ldots, p\}^* \to \{0, \ldots, p-1\}^*$$

which gives the normalized representation in base p of an integer (the representation obtained by the greedy algorithm) is a rational function, its graph $\hat{\nu}$ is recognizable by a finite letter-to-letter automaton. $\square$

Lemma 15 [5] *A subset of $\mathbb{N}$ is recognizable in base $p \geq 2$ if and only if it is definable in the structure $\langle \mathbb{N}, +, V_p \rangle$, where for $x \neq 0$, $V_p(x)$ is the greatest power of p dividing x while $V_p(0) = 1$.* $\square$

Proposition 16 *Let $S = (L, \Sigma, <)$ be a numeration system. For each natural number t, $X + t$ is S-recognizable if $X \subset \mathbb{N}$ is S-recognizable.*

Proof. Let $\Sigma = \{\sigma_1 < \cdots < \sigma_p\}$ and let the homomorphism $h : \Sigma^* \to \{1, \ldots, p\}^*$ be defined by $h : \sigma_i \mapsto i$. For $x \in \mathbb{N}$, the word $h(\mathrm{r}_S(x)) = x_0 \ldots x_l \in \{1, \ldots, p\}^*$ is a representation in base p of the integer $\pi_p(h(\mathrm{r}_S(x))) = x_0\, p^l + \cdots + x_l\, p^0$.

Since L is regular over Σ, by Lemma 14, $\nu(h(L))$ is regular over $\{0, \ldots, p-1\}$ and by Lemma 15, the set

$$\mathcal{N} = \pi_p[\nu(h(L))]$$

is definable in $\langle \mathbb{N}, +, V_p \rangle$ (we say that $\mathcal{N}$ is p-definable). In other words, $\pi_p \circ \nu \circ h$ is a one-to-one correspondence which maps regular subsets of L onto p-definable subsets of $\mathbb{N}$.

The successor function $S_L : L \to L$ (with respect to the lexicographic order) is then regular (i.e., maps regular subsets onto regular subsets). Indeed, $\mathcal{S} = \pi_p \circ \nu \circ h \circ S_L \circ (\pi_p \circ \nu \circ h)^{-1}$ is the restriction to $\mathcal{N}$ of the function $x \mapsto y$ defined in $\langle \mathbb{N}, +, V_p \rangle$ by the formula

$$(y \in \mathcal{N}) \wedge (x < y) \wedge (\forall z)(z \in \mathcal{N} \wedge x < z) \to (y \leq z).$$

So, $\mathcal{S}$ maps p-definable sets onto p-definable sets. The commutative diagram on Figure 3 summarizes the situation,

$$\begin{tikzcd}
\mathrm{r}_S(\mathbb{N}) = L \arrow[r, "\pi_p \circ \nu \circ h"] \arrow[d, "S_L"'] & \mathcal{N} \subset \mathbb{N} \arrow[d, "\mathcal{S}"] \\
\mathrm{r}_S(\mathbb{N} \setminus \{0\}) \subset L & \mathcal{S}(\mathcal{N}) \subset \mathcal{N} \arrow[l, "(\pi_p \circ \nu \circ h)^{-1}"]
\end{tikzcd}$$

Figure 3: The different applications.

Assume now that X is S-recognizable, i.e., that $\mathrm{r}_S(X)$ is a regular set. Then $\mathrm{r}_S(X + t) = S_L^t(\mathrm{r}_S(X))$ is regular. $\square$

6 Multiplication by a constant

In this section, we show that, in general, multiplication by a constant does not preserve recognizability. To that end, we use the system $S = (a^*b^*, \{a, b\}, a < b)$, for which it is easy to see that

$$\mathrm{val}_S(a^p b^q) = \frac{1}{2}(p + q)(p + q + 1) + q.$$

Remark 3 Observe that the r.h.s. is the well-known function of Peano [25].

It would suffice to show that, say, multiplication by two does not preserve recognizability but here we are lucky enough to get more.

Theorem 17 *Let S be the numeration system $(a^*b^*, \{a, b\}, a < b)$ and let $\alpha \in \mathbb{N}$. Multiplication by α transforms the S-recognizable sets into S-recognizable sets if and only if α is an odd perfect square.*

Proof. (i)*Sketch.* If α is not a perfect square, we show that for a suitably chosen r,

$$\mathcal{L}_\alpha^r = a^r b^* \cap \mathrm{r}_S(\alpha \mathrm{val}_S(a^*))$$

is infinite while the set of lengths $|\mathcal{L}_\alpha^r|$ only contains finite arithmetic progressions so that $\mathrm{r}_S(\alpha\mathrm{val}_S(a^*))$ is not even context-free, thanks to Parikh's theorem [17].

If $\alpha = \beta^2$, $\mathbb{N}^2$ is divided into $\beta+1$ regions R_i in each of which an explicit formula for the function $M : (p,q) \mapsto (r,s)$ such that $\alpha\,\mathrm{val}_S(a^p b^q) = \mathrm{val}_S(a^r b^s)$ can be supplied. These regions come from length considerations: given a word of length l and of numerical value x, there are $\beta+1$ possible lengths for the word of value αx. When α is an odd perfect square, the fact that multiplication by α preserves the regularity of the subsets of a^*b^* then follows from an easy lemma. If $\alpha = (2\gamma)^2$ then one shows that there exist constants j and k such that for n large enough, $\mathrm{r}_S(\alpha\,\mathrm{val}_S(a^n)) = a^{j+n\gamma}b^{k+n\gamma}$.

(ii) *Case of a non-perfect-square.* Let α be a non-perfect-square integer. We have

$$l \in |\mathcal{L}_\alpha^r| \Leftrightarrow \exists p : \ \mathrm{val}_S(a^r b^{l-r}) = \alpha\,\mathrm{val}_S(a^p).$$

In other words, $l \in |\mathcal{L}_\alpha^r|$ if and only if

$$[2(r+s)+3]^2 - \alpha(2p+1)^2 = 8r+9-\alpha \tag{2}$$

for some p, where $s = l - r$.

To guarantee that $|\mathcal{L}_\alpha^r|$ be infinite, we choose r in such a way that

$$X^2 - \alpha Y^2 = 8r + 9 - \alpha \tag{3}$$

has infinitely many solutions (X,Y) with X, Y odd. To that end, it suffices to choose r such that $8r+9-\alpha > 0$ and that the equation (3) admits a solution $(x,1)$ with x odd (cf. Appendix). This can be achieved with r of the form z^2. Indeed, the equation $x^2 - 8z^2 = 9$ has infinitely many solutions given by

$$\begin{pmatrix} x_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \ \begin{pmatrix} x_{i+1} \\ z_{i+1} \end{pmatrix} = \begin{pmatrix} 3 & 8 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_i \\ z_i \end{pmatrix}, \ \forall i \in \mathbb{N}.$$

The x_i's are odd. We choose i such that $8z_i^2 + 9 - \alpha > 0$ and take $x = x_i$.

The set of solutions of (3) with odd components is a finite union of sequences $(X_n^{(j)}, Y_n^{(j)})_{n\in\mathbb{N}}$, $j = 1,\ldots,m$, such that $X_n^{(j)} > C^n$ for some $C > 1$ (cf. Appendix).

We are now in a position to show that $|\mathcal{L}_\alpha^r|$ only contains finite arithmetic progressions. Suppose to the contrary that it contains an infinite progression. Then there exist $\lambda, \mu \in \mathbb{N}, \mu > 0$, and, for each $t \in \mathbb{N}$, indices $n_t \in \mathbb{N}, j_t \in \{1,\ldots,m\}$ such that

$$\lambda + \mu t = X_{n_t}^{(j_t)} > C^{n_t}.$$

Given t, the sequence $n_0, \ldots, n_{mt}$ contains at least t distinct numbers. Therefore

$$\forall t \in \mathbb{N},\ \lambda + \mu m t > C^t,$$

a contradiction.

(iii) *The case of an odd perfect square.* Let $\alpha = \beta^2$ and β be an odd integer.

We want to compute r, s such that $\alpha \,\mathrm{val}_S(a^p b^q) = \mathrm{val}_S(a^r b^s)$, i.e.,

$$[2(r+s)+3]^2 - \beta^2[2(p+q)+3]^2 = 8r - 8p\beta^2 - 9(\beta^2 - 1).$$

Let $l = p+q$, $l' = r+s$. Then

$$\alpha l(l+1) \leq 2\alpha \mathrm{val}_S(a^p b^q) \leq \alpha l(l+3) \text{ and } l'(l'+1) \leq 2\mathrm{val}_S(a^r b^s) \leq l'(l'+3).$$

Therefore, $l'(l'+1) \leq \beta^2 l(l+3)$ and $\beta^2 l(l+1) \leq l'(l'+3)$. From this, it follows easily that

$$r + s = \beta(p+q) + \left\lfloor \frac{\beta}{2} \right\rfloor + i$$

and thus

$$\begin{cases} r = r_i(p,q) := \beta(i+1)p - \beta(\beta - i - 1)q + \frac{1}{8}[(\beta + 2i + 2)^2 - 9] \\ s = s_i(p,q) := -\beta i p + \beta(\beta - i)q - \frac{1}{8}[(\beta + 2i)^2 - 9] - 1 \end{cases}$$

for some $i \in \{-1, \ldots, \beta - 1\}$. These equations together with the conditions $r, s \geq 0$ define $\beta + 1$ regions R_i which divide $\mathbb{N}^2$.

The regular subsets of $a^* b^*$ are the finite unions of sets of the form

$$D = \{a^{y+fz} b^{w+gx} : f, g \geq 0\},$$

$w, x, y, z \geq 0$. Substituting $y + fz$ and $w + gx$ in place of p and q, respectively, in $r_i(p,q)$ and $s_i(p,q)$, one sees that $D' = \mathrm{r}_S[\alpha \,\mathrm{val}_S(D \cap R_i)]$ is of the form (4) of Lemma 18 below, the matrix A being

$$A = \begin{pmatrix} z\beta(i+1) & -x\beta(\beta - i - 1) \\ -z\beta i & x\beta(\beta - i) \end{pmatrix}.$$

One can apply the lemma to see that D' is regular except if $i = -1$ or $xz = 0$. In these cases, D' is easily shown to be regular by direct inspection.

(iv) *The case of an even perfect square.* Let $\alpha = \beta^2$ with $\beta = 2\gamma$, $\gamma \in \mathbb{N} \setminus \{0\}$. Using the same kind of computation as in (iii), one has, for p large enough, $\mathrm{r}_S(\alpha \,\mathrm{val}_S(a^p)) = a^r b^s$ with

$$\begin{cases} r &= \gamma p + \frac{1}{2}\gamma(\gamma + 1) - 1 \\ s &= \gamma p - \frac{1}{2}\gamma(\gamma - 1). \end{cases}$$

Therefore, $\mathrm{r}_S(\alpha \,\mathrm{val}_S(a^*))$ is context-free but not regular. □

Lemma 18 *Let A be a non-singular $p \times p$ integral matrix. For $i = 1, \ldots, p$, set*

$$h_i(\mathbf{n}) = A_{i1}n_1 + \cdots + A_{ip}n_p - b_i,$$

where $\mathbf{n} = (n_1, \ldots, n_p) \in \mathbb{N}^p$ and $b_1, \ldots, b_p \in \mathbb{Z}$. If the entries of $\det(A)A^{-1}$ are non-negative, then the language

$$\mathcal{L} = \{a_1^{h_1} \cdots a_p^{h_p} : h_1(\mathbf{n}) \geq 0, \ldots, h_p(\mathbf{n}) \geq 0, \mathbf{n} \in \mathbb{N}^p\} \tag{4}$$

is a regular subset of $a_1^* \cdots a_p^*$.

Proof. If $\mathbf{n} \in \mathbb{N}^p$ satisfies $h_i(\mathbf{n}) \geq 0$ then $(A\mathbf{n})_i = b_i + u_i$, i.e.,

$$n_i = \sum_{j=1}^{p} (A^{-1})_{ij}(b_j + u_j), \tag{5}$$

for some $u_i \in \mathbb{N}$.

We need to describe those $\mathbf{u} = (u_1, \ldots, u_p) \in \mathbb{N}^p$ for which (5) defines non-negative integers n_i.

If $\det(A) < 0$, the entries of A^{-1} are negative, there are finitely many such $\mathbf{u}$ and $\mathcal{L}$ is finite. If $\det(A) > 0$, $(A^{-1})_{ij} \geq 0$, for large enough u_j's, (5) defines thus positive numbers n_i but it remains to ensure that they are integers. To that purpose, since $A^{-1} = \mathcal{A}/\det(A)$, where the entries of $\mathcal{A}$ are natural numbers, it is necessary and sufficient that the remainders $r_j \in \{0, \ldots, \det(A) - 1\}$ of the division of u_j by $\det(A)$ satisfy

$$\sum_{j=1}^{p} \mathcal{A}_{ij}(b_j + r_j) \equiv 0 \pmod{\det(A)}.$$

There is a finite number of such $(r_1, \ldots, r_p)$ so that $\mathcal{L}$ is a finite union of regular languages of the form

$$\left(a_1^{\det(A)}\right)^* a_1^{s_1\det(A)+r_1} \ldots \left(a_p^{\det(A)}\right)^* a_p^{s_p\det(A)+r_p}.$$

(The s_j's are chosen to guarantee that the u_j's be large enough for the corresponding n_i's to be non-negative.) □

7 Appendix

a) The next proposition summarizes some well-known facts that are used in the proof of Theorem 17. The reader will find in [8, 24] the material necessary to achieve its proof.

Proposition 19 *Assume that $\alpha \in \mathbb{N}$ is not a perfect square and that $N > 0$ is a natural number.*
(i) The set of solutions $(X, Y) \in \mathbb{N}^2$ of the equation $X^2 - \alpha Y^2 = N$ is the (finite) union of the sequences $(X_n, Y_n)_{n \in \mathbb{N}}$ defined by

$$\begin{pmatrix} X_{i+1} \\ Y_{i+1} \end{pmatrix} = \begin{pmatrix} u & \alpha v \\ v & u \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \end{pmatrix}, \ \forall i \in \mathbb{N}, \text{ and } 0 < X_0 \leq u\sqrt{N}, \qquad (6)$$

where $(u, v) \in \mathbb{N}^2$ is the minimal non-trivial solution of $U^2 - \alpha V^2 = 1$, i.e., that for which $u > 1$ is the smallest.
(ii) Each component of any solution $(X_n, Y_n)_{n \in \mathbb{N}}$ of (6) are solutions of

$$Z_{i+2} = 2uZ_{i+1} - Z_i, \ \forall i \in \mathbb{N}.$$

In particular, X_{2n}, X_{2n+1}, Y_{2n} and Y_{2n+1} are of the same parity as X_0, X_1, Y_0 and Y_1 respectively.
(iii) For any solution $(X_n, Y_n)_{n \in \mathbb{N}}$ of (6), one has $X_n > u^n$. □

b) Taking advantage of Lemmas 14 and 15, we give another proof of Theorem 4, based on the notion of substitution.

Lemma 20 [4, 5, 7] *A subset X of $\mathbb{N}$ is recognizable in base p if and only if the characteristic sequence of X is generated by a p-substitution.* □

Lemma 21 [7] *The set of the infinite words generated by p-substitution is closed under finite transduction.* □

Proof of Theorem 4. We use the notation of Proposition 16. The set $\nu(h(L))$ is a regular subset of $\{0, \ldots, |\Sigma| - 1\}^*$ and by Lemma 20, the characteristic sequence Ψ of $\pi_{|\Sigma|}[\nu(h(L))]$ is generated by a $|\Sigma|$-substitution. To conclude, use Lemma 21 and observe that the characteristic sequence of $\pi_{|\Sigma|}[\nu(h(p + \mathbb{N}q))]$ is the image of Ψ under the following finite transducer on Figure 4 (the tail has p nodes and the head counts q of them) each state has a loop which

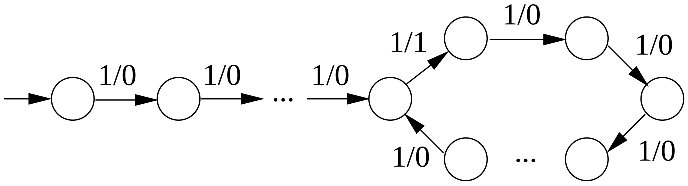

Figure 4: The finite transducer for $\pi_{|\Sigma|}[\nu(h(p + \mathbb{N}q))]$

corresponds to the reading and the writing of 0. □

c) The nature of the S-recognizable sets seems to depend strongly on the system S. In standard numeration systems with an integer base, the set of squares is not recognizable [11, p. 311] while an example of a system for which it is recognizable may be found in [19, p. 141]. Here is another example, based on Lemma 9.

Proposition 22 *Let $S = (a^*b^* \cup a^*c^*, \{a, b, c\}, a < b < c)$. The set $\{n^2 : n \in \mathbb{N}\}$ is S-recognizable.*

Proof. Indeed, since $\#((a^*b^* \cup a^*c^*) \cap \Sigma^n) = 2n + 1$, the greatest word of length n in $a^*b^* \cup a^*c^*$ has numerical value n^2. □

Using the same idea, one can easily produce various examples of unusual recognizable sets, such as $\{v_n : n \in \mathbb{N}\}$ for any regular language L.

Acknowledgments

We would like to thank V. Bruyère for her valuable advice and encouragements and M. De Wilde for fruitful comments. We are grateful to the referees for valuable advice in improving the presentation of this paper.

References

[1] M. Andraşiu, J. Dassow, G. Păun, A. Salomaa, Language-theoretic problems arising from Richelieu cryptosystems, *Theor. Comput. Sci.* **116** (1993), 339–357.

[2] J. Berstel, C. Reutenauer, *Les séries rationnelles et leurs langages*, Études et Recherches en Informatique, Masson, Paris, 1984.

[3] A. Bertrand-Mathis, Comment écrire les nombres entiers dans une base qui n'est pas entière, *Acta Math. Acad. Sci. Hungar.* **54** (1989), 237–241.

[4] V. Bruyère, G. Hansel, Bertrand numeration systems and recognizability, *Theor. Comput. Sci.* **181** (1997), 17–43.

[5] V. Bruyère, G. Hansel, C. Michaux, R. Villemaire, Logic and p-recognizable sets of integers, *Bull. Belg. Math. Soc.* **1** (1994), 191–238.

[6] A. Cobham, On the base-dependence of sets of numbers recognizable by finite automata, *Math. Systems Theory* **3** (1969), 186–192.

[7] A. Cobham, Uniform tag sequences, *Math. Systems Theory* **6** (1972), 186–192.

[8] H. Davenport, *The Higher Arithmetic*, Cambridge University Press, Sixth Ed., 1992.

[9] F. Durand, A generalization of Cobham's theorem, *Theory Comput. Syst.* **31** (1998), 169–185.

[10] F. Durand, Sur les ensembles d'entiers reconnaissables, *J. Théor. Nombres Bordeaux* **10** (1998), 65–84.

[11] S. Eilenberg, *Automata, Languages and Machines*, Vol. A, Academic Press, New York, 1974.

[12] A. S. Fraenkel, Systems of numeration, *Amer. Math. Monthly* **92** (1985), 105–114.

[13] C. Frougny, Representations of numbers and finite automata, *Math. Systems Theory* **25** (1992), 37–60.

[14] C. Frougny, B. Solomyak, On representation of integers in linear numeration systems, in *Ergodic theory of Z^d actions* (Warwick, 1993–1994), 345–368, London Math. Soc. Lecture Note Ser. **228**, Cambridge Univ. Press, Cambridge, 1996.

[15] G. Hansel, Systèmes de numération indépendants et syndéticité, *Theor. Comput. Sci.* **204** (1998), 119–130.

[16] M. Hollander, Greedy numeration systems and regularity, *Theory Comput. Syst.* **31** (1998), 111–133.

[17] H. R. Lewis, C. H. Papadimitriou, *Elements of the Theory of Computation*, Prentice-Hall, 1981.

[18] N. Loraud, β-shift, systèmes de numération et automates, *J. Théor. Nombres Bordeaux* **7** (1995), 473–498.

[19] A. Maes, Morphic predicates and applications to the decidability of arithmetic theories, *Thesis*, Université de Mons-Hainaut (1999).

[20] G. Păun, A. Salomaa, Thin and slender languages, *Discrete Appl. Math.* **61** (1995), 257–270.

[21] D. Perrin, Finite automata, in *Handbook of Theoretical Computer Science*, Vol. B, J. Van Leeuwen, Ed., Elsevier, Amsterdam, 1990, 2–57.

[22] F. Point and V. Bruyère, On the Cobham-Semenov theorem, *Theory Comput. Syst.* **30** (1997), 197–220.

[23] J. Shallit, Numeration systems, linear recurrences, and regular sets, *Information and Computation*, **113** (1994), 331–347.

[24] H. E. Rose, *A Course in Number Theory*, Oxford Science Pub., Second Ed., 1996.

[25] S. Yablonski, *Introduction aux mathématiques discrètes*, Ed. Mir, Moscou, 1983.