RAIRO-Theor. Inf. Appl. 46 (2012) 51–65
DOI: 10.1051/ita/2011112

Available online at:
www.rairo-ita.org

MULTI-DIMENSIONAL SETS RECOGNIZABLE IN ALL ABSTRACT NUMERATION SYSTEMS

Émilie Charlier[1], Anne Lacroix[1] and Narad Rampersad[1]

Abstract. We prove that the subsets of $\mathbb{N}^d$ that are S-recognizable for all abstract numeration systems S are exactly the 1-recognizable sets. This generalizes a result of Lecomte and Rigo in the one-dimensional setting.

Mathematics Subject Classification. 68Q45.

1. Introduction

In this paper we characterize the subsets of $\mathbb{N}^d$ that are simultaneously recognizable in all abstract numeration systems (numeration systems that represent a natural number n by the $(n+1)$-th word of a genealogically ordered regular language – see below for the precise definition). Lecomte and Rigo [11] provided such a characterization for the case $d = 1$ based on the well-known correspondence between unary regular languages and ultimately periodic subsets of $\mathbb{N}$. When $d > 1$ we no longer have such a nice correspondence and the situation becomes somewhat more complicated. To obtain our characterization we instead use a classical decomposition theorem due to Eilenberg *et al.* [7]. The motivation for studying such sets comes from the well-known result of Cobham (and its multi-dimensional generalization due to Semenov) concerning the sets recognizable in integer bases.

Let $k \geq 2$ be an integer. A set $X \subseteq \mathbb{N}$ is *k-recognizable* (or *k-automatic*) if the language consisting of the base-k representations of the elements of X is accepted by a finite automaton. A celebrated result of Cobham [5] characterizes the sets that are recognizable in all integer bases $k \geq 2$. Two numbers k and ℓ are *multiplicatively independent* if $k^m = \ell^n$ implies $m = n = 0$. A subset of the integers is *ultimately periodic* if it is a finite union of arithmetic progressions.

Keywords and phrases. Finite automata, numeration systems, recognizable sets of integers, multi-dimensional setting.

[1] Department of Mathematics, University of Liège, Grande Traverse 12 (B37), 4000 Liège, Belgium. {echarlier,A.Lacroix,nrampersad}@ulg.ac.be

Theorem 1.1 (Cobham)**.** *Let $k, \ell \geq 2$ be two multiplicatively independent integers and let $X \subseteq \mathbb{N}$. The set X is both k-recognizable and ℓ-recognizable if and only if it is ultimately periodic.*

We say that a set $X \subseteq \mathbb{N}$ is 1-*recognizable* if the language $\{a^n : n \in X\}$ consisting of the unary representations of the elements of X is accepted by a finite automaton. It is well-known [6], Proposition V.1.1, that a set is 1-recognizable if and only if it is ultimately periodic.

Lecomte and Rigo [11] introduced the following generalization of the standard integer base numeration systems.

Definition 1.2. An *abstract numeration system* is a triple $S = (L, \Sigma, <)$ where L is an infinite regular language over a totally ordered finite alphabet $(\Sigma, <)$. The map $\mathrm{rep}_S : \mathbb{N} \to L$ is a bijection mapping $n \in \mathbb{N}$ to the $(n+1)$-th word of L ordered genealogically. The inverse map is denoted by $\mathrm{val}_S : L \to \mathbb{N}$. A set $X \subseteq \mathbb{N}$ is S-*recognizable* if the language $\mathrm{rep}_S(X) = \{\mathrm{rep}_S(n) : n \in X\}$ is regular.

Lecomte and Rigo [11] proved that any ultimately periodic set is S-recognizable for any abstract numeration system S. Suppose on the other hand that $X \subseteq \mathbb{N}$ is S-recognizable for every abstract numeration system S. Then in particular, the set X must be 1-recognizable, and hence must be ultimately periodic. We therefore have the following characterization of the sets that are recognizable in all abstract numeration systems.

Theorem 1.3 (Lecomte and Rigo)**.** *A set $X \subseteq \mathbb{N}$ is S-recognizable for all abstract numeration systems S if and only if it is ultimately periodic.*

Rigo and Maes [14] considered S-recognizability in a multi-dimensional setting. This concept was further studied by Charlier *et al.* [4]. For the formal definitions we need to introduce the following "padding" function.

Definition 1.4. If $w_1, \ldots, w_d$ are finite words over the alphabet Σ, the padding map

$$(\cdot)^{\#} : (\Sigma^*)^d \to ((\Sigma \cup \{\#\})^d)^*$$

is defined by

$$(w_1, \ldots, w_d)^{\#} := (w_1 \#^{m-|w_1|}, \ldots, w_d \#^{m-|w_d|})$$

where $m = \max\{|w_1|, \ldots, |w_d|\}$. Here we write (ac, bd) to denote the concatenation $(a, b)(c, d)$.

If $R \subseteq (\Sigma^*)^d$, then

$$R^{\#} = \{(w_1, \ldots, w_d)^{\#} : (w_1, \ldots, w_d) \in R\}.$$

Note that R is not necessarily a language, whereas $R^{\#}$ is; that is, the set R consists of d-tuples of words over Σ, whereas $R^{\#}$ consists of words over the alphabet $(\Sigma \cup \{\#\})^d$.

Definition 1.5. Let $S = (L, \Sigma, <)$ be an abstract numeration system. Let $X \subseteq \mathbb{N}^d$. The set X is *S-recognizable* (or *S-automatic*) if the language $\mathrm{rep}_S(X)^{\#}$ is regular, where

$$\mathrm{rep}_S(X) = \{(\mathrm{rep}_S(n_1), \ldots, \mathrm{rep}_S(n_d)) \colon (n_1, \ldots, n_d) \in X\}.$$

Let $k \geq 2$ be an integer. The notions of k-recognizability and 1-recognizability are special cases of S-recognizability. The set X is *k-recognizable* (or *k-automatic*) if it is S-recognizable for the abstract numeration system S built on the language consisting of the base-k representations of the elements of X. The set X is *1-recognizable* (or *1-automatic*) if it is S-recognizable for the abstract numeration system S built on a^*.

The multi-dimensional analogue of Cobham's theorem, due to Semenov [16], requires an analogous notion of ultimate periodicity in the multi-dimensional setting.

Definition 1.6. A set $X \subseteq \mathbb{N}^d$ is *linear* if there exists $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_t \in \mathbb{N}^d$ such that

$$X = \{\mathbf{v}_0 + n_1\mathbf{v}_1 + n_2\mathbf{v}_2 + \ldots + n_t\mathbf{v}_t : n_1, \ldots, n_t \in \mathbb{N}\}.$$

A set $X \subseteq \mathbb{N}^d$ is *semi-linear* if it is a finite union of linear sets.

For more on semi-linear sets see [10]. We can now state the multi-dimensional version of Cobham's Theorem [16].

Theorem 1.7 (Cobham-Semenov)**.** *Let $k, \ell \geq 2$ be two multiplicatively independent integers and let $X \subseteq \mathbb{N}^d$. The set X is both k-recognizable and ℓ-recognizable if and only if it is semi-linear.*

In other words, the semi-linear sets are precisely the sets recognizable in all integer bases $k \geq 2$. One might therefore expect that, as in Theorem 1.3, the semi-linear sets are recognizable in all abstract numeration systems. However, this fails to be the case, as the following example shows.

Example 1.8. The semi-linear set $X = \{n(1, 2) : n \in \mathbb{N}\} = \{(n, 2n) : n \in \mathbb{N}\}$ is not 1-recognizable. Consider the language $\{(a^n\#^n, a^{2n}) : n \in \mathbb{N}\}$, consisting of the unary representations of the elements of X. An easy application of the pumping lemma shows that this is not a regular language.

Observe that in the one-dimensional case, we have the following equivalences: semi-linear $\Leftrightarrow$ ultimately periodic $\Leftrightarrow$ 1-recognizable. However, Example 1.8 shows that these equivalences no longer hold in the multi-dimensional setting. In order to get a multi-dimensional analogue of Theorem 1.3, we must consider the class of 1-recognizable sets, which form a proper subclass of the class of semi-linear sets.

Another well-studied subclass of the class of semi-linear sets is the class of recognizable sets. A subset X of $\mathbb{N}^d$ is *recognizable* if there exists a finite monoid M, a monoid homomorphism $\varphi : \mathbb{N}^d \to M$, and a subset $B \subseteq M$ such that $X = \varphi^{-1}(B)$. When $d = 1$, we have again the following equivalences: recognizable

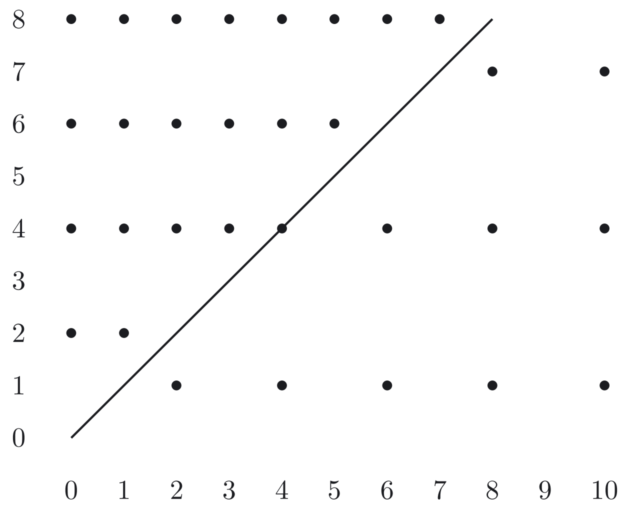

FIGURE 1. The set X of Example 1.10.

$\Leftrightarrow$ ultimately periodic $\Leftrightarrow$ 1-recognizable. However, for $d > 1$ these equivalences no longer hold. An unpublished result of Mezei (see [6], Prop. III.12.2) demonstrates that the recognizable subsets of $\mathbb{N}^2$ are precisely finite unions of sets of the form $Y \times Z$, where Y and Z are ultimately periodic subsets of $\mathbb{N}$. In particular, the *diagonal set* $D = \{(n, n) : n \in \mathbb{N}\}$ is not recognizable [6], Exercise III.12.7. However, the set D is clearly a 1-recognizable subset of $\mathbb{N}^2$. So we see that for $d > 1$, the class of 1-recognizable sets corresponds neither to the class of semi-linear sets, nor to the class of recognizable sets. For further information on recognizable sets, their different characterizations and the classical Cobham-Semenov Theorem, see [3].

Our main result is the following, which generalizes the result of Lecomte and Rigo (Thm. 1.3).

Theorem 1.9. *Let $X \subseteq \mathbb{N}^d$. Then X is S-recognizable for all abstract numeration systems S if and only if X is 1-recognizable.*

To illustrate this theorem, we give the following example.

Example 1.10. Let

$$X = \{(2n, 3m + 1) : n, m \in \mathbb{N} \text{ and } 2n \geq 3m + 1\} \quad \cup \\ \{(n, 2m) : n, m \in \mathbb{N} \text{ and } n < 2m\}$$

(see Fig. 1). It is clear that X is 1-recognizable. Let $S = (L, \Sigma, <)$ be an abstract numeration system. By Theorem 1.3, the sets $\{2n : n \in \mathbb{N}\}$ and $\{3m+1 : m \in \mathbb{N}\}$ are both S-recognizable, and so the set $\{(2n, 3m + 1) : n, m \in \mathbb{N}\}$ is also S-recognizable. In other words, the set $\{(\text{rep}_S(2n), \text{rep}_S(3m+1))^{\#} : n, m \in \mathbb{N}\}$ is accepted by a finite automaton. Furthermore, the set $\{(x, y)^{\#} : x, y \in L \text{ and } x \geq y\}$

is also accepted by a finite automaton, and so by taking the product of these two automata we obtain an automaton accepting

$$\{(\mathrm{rep}_S(2n), \mathrm{rep}_S(3m+1))^\# : n, m \in \mathbb{N} \text{ and } 2n \geq 3m+1\}.$$

In the same way we can construct an automaton to accept the set

$$\{(\mathrm{rep}_S(n), \mathrm{rep}_S(2m))^\# : n, m \in \mathbb{N} \text{ and } n < 2m\}.$$

Since the union of two regular languages is regular, we see that X is S-recognizable.

2. Proof of our main result

In order to obtain our main result, we will need a classical result of Eilenberg *et al.* [7], Theorem 11.1 (see also [15], Thm. C.1.1). We first need the following definition.

Definition 2.1. Let A be a non-empty subset of $\{1, \dots, d\}$. Define the subalphabet

$$\Sigma_A = \{x \in (\Sigma \cup \{\#\})^d : \text{the } i\text{-th component of } x \text{ is } \# \text{ exactly when } i \notin A\}.$$

Example 2.2. Let $\Sigma = \{a, b\}$ and $d = 4$. If $A = \{1, 2, 3, 4\}$, then $\Sigma_A = \{(\sigma_1, \sigma_2, \sigma_3, \sigma_4) : \sigma_i \in \Sigma \text{ for } i = 1, 2, 3, 4\}$. If $A = \{2, 3\}$, then $\Sigma_A = \{(\#, \sigma_2, \sigma_3, \#) : \sigma_i \in \Sigma \text{ for } i \in \{2, 3\}\}$. If $A = \{3\}$, then $\Sigma_A = \{(\#, \#, a, \#), (\#, \#, b, \#)\}$.

Theorem 2.3 (Decomposition [7])**.** *Let $R \subseteq (\Sigma^*)^d$. The language $R^\# \subseteq ((\Sigma \cup \{\#\})^d)^*$ is regular if and only if it is a finite union of languages of the form*

$$R_0 \dots R_t, \quad t \in \mathbb{N},$$

where each factor $R_i \subseteq (\Sigma_{A_i})^*$ is regular and $A_t \subseteq \dots \subseteq A_0 \subseteq \{1, \dots, d\}$.

Remark 2.4. Theorem 2.3 does not hold if $R^\#$ is replaced by an arbitrary language over $(\Sigma \cup \{\#\})^d$. It is only valid due to the definition of the map $(\cdot)^\#$.

Example 2.5. Let $R = \{(a^{5n}, a^{6m}) : n, m \in \mathbb{N}\}$. Then $R^\#$ is regular, since one can easily construct an automaton that simultaneously checks that the length of the first component of its input is a multiple of 5 and that the length of the second component is a multiple of 6. Moreover, we have

$$R^\# = \bigcup_{\ell=0}^{5} (a^{30}, a^{30})^* (a^{5\ell}\#^\ell, a^{6\ell})(\#^6, a^6)^* \quad \cup \\ \bigcup_{\ell=0}^{4} (a^{30}, a^{30})^* (a^{5(\ell+1)}, a^{6\ell}\#^{5-\ell})(a^5, \#^5)^*.$$

Observe that each of the languages appearing in the unions above are products of the form described in Theorem 2.3.

Lemma 2.6. *Let $X \subseteq \mathbb{N}^d$. Then X is 1-recognizable if and only if X is a finite union of sets of the form*

$$\left\{ \sum_{\ell=0}^{t} (c_\ell(n_{\ell,1}, \dots, n_{\ell,d}) + (b_{\ell,1}, \dots, b_{\ell,d}))\colon (\forall \ell)(\forall i)\ n_{\ell,i} \in \mathbb{N} \ \text{ and } \right. \\ \left. (\forall \ell)(\forall i)\ (i \notin A_\ell \Rightarrow n_{\ell,i} = 0) \ \text{and}\ (\forall \ell)(\forall i)(\forall j)(i, j \in A_\ell \Rightarrow n_{\ell,i} = n_{\ell,j}) \right\} \tag{2.1}$$

where

- $t \in \mathbb{N}$;
- $A_t \subseteq \ldots \subseteq A_0 \subseteq \{1, \dots, d\}$;
- $c_0, \dots, c_t \in \mathbb{N}$;
- $(\forall \ell)(\forall i)\ b_{\ell,i} \in \mathbb{N}$;
- $(\forall \ell)(\forall i)\ (i \notin A_\ell \Rightarrow b_{\ell,i} = 0)$; *and*
- $(\forall \ell)(\forall i)(\forall j)\ (i, j \in A_\ell \Rightarrow b_{\ell,i} = b_{\ell,j})$.

Proof. Let $\Sigma = \{a\}$ and let $S = (\Sigma^*, \Sigma, <)$. We define

$$R := \text{rep}_S(X) = \{(a^{n_1}, \dots, a^{n_d})\colon (n_1, \dots, n_d) \in X\}.$$

The set X is 1-recognizable if and only if the language $R^\#$ is regular. By Theorem 2.3, the language $R^\#$ is regular if and only if it is a finite union of languages of the form

$$R_0 \dots R_t, \quad t \in \mathbb{N},$$

where each factor $R_\ell \subseteq (\Sigma_{A_\ell})^*$ is regular and $A_t \subseteq \ldots \subseteq A_0 \subseteq \{1, \dots, d\}$. Since $|\Sigma| = 1$, we have $|\Sigma_{A_\ell}| = 1$. Let $\Sigma_{A_\ell} = \{x\}$. It is well-known [6], Prop. V.1.1, that R_ℓ is a finite union of languages of the form $\{x^{pi+q} : i \in \mathbb{N}\}$, where $p, q \in \mathbb{N}$. Without loss of generality we can assume that R_ℓ is exactly of this form. Hence, the language R_ℓ consists of the representations of a set of the form

$$\{c_\ell(n_{\ell,1}, \dots, n_{\ell,d}) + (b_{\ell,1}, \dots, b_{\ell,d}) : (\forall i)(n_{\ell,i} \in \mathbb{N})\}.$$

The conditions $A_t \subseteq \ldots \subseteq A_0 \subseteq \{1, \dots, d\}$ impose the restrictions on the $n_{\ell,i}$'s and the constants $b_{\ell,i}$ in the statement of the lemma. The concatenation of the R_ℓ's gives the sum described above. □

Remark 2.7. We can give an alternative description of the 1-recognizable sets. Let $\mathbf{v} = (v_1, \dots, v_d) \in \mathbb{N}^d$. We define $\text{Supp}(\mathbf{v}) = \{i \in \{1, \dots, d\} : v_i \neq 0\}$. Let $X \subseteq \mathbb{N}^d$. Then X is a finite union of sets of the form described in Lemma 2.6 if and only if X is a finite union of sets of the form

$$(b_0 + c_0\,\mathbb{N})\mathbf{v_0} + \ldots + (b_t + c_t\,\mathbb{N})\mathbf{v_t},$$

where

- $t \in \mathbb{N}$;
- $b_i, c_i \in \mathbb{N}$ for $i = 1, \ldots, t$;
- $\mathbf{v}_i \in \{0,1\}^d$ for $i = 1, \ldots, t$;
- $\mathrm{Supp}(\mathbf{v}_t) \subseteq \ldots \subseteq \mathrm{Supp}(\mathbf{v}_0)$.

Example 2.8. Let $X = \{(5n, 5n+4m+6\ell+1, 5n+4m+6\ell+3, 5n) : n, m, \ell \in \mathbb{N}\}$. The unary representation of X is

$$R^{\#} = ((a,a,a,a)^5)^*((\#,a,a,\#)^4)^*((\#,a,a,\#)^6)^*(\#,a,a,\#)(\#,\#,a,\#)^2.$$

Since $R^{\#}$ is regular the set X is 1-recognizable. The set X can be written as

$$X = \{5(n,n,n,n)+4(0,m,m,0)+6(0,\ell,\ell,0)+(0,1,1,0)+(0,0,2,0) : n,m,\ell \in \mathbb{N}\},$$

which is an expression of the form (2.1) where $t = 3$; $A_0 = \{1,2,3,4\}$, $A_1 = A_2 = \{2,3\}$, $A_3 = \{3\}$; $c_0 = 5, c_1 = 4, c_2 = 6, c_3 = 0$; and $b_{0,i} = b_{1,i} = 0$ for all i, $(b_{2,1}, b_{2,2}, b_{2,3}, b_{2,4}) = (0,1,1,0)$, $(b_{3,1}, b_{3,2}, b_{3,3}, b_{3,4}) = (0,0,2,0)$.

Alternatively, by Remark 2.7 we can write

$$X = 5\,\mathbb{N}(1,1,1,1) + 4\,\mathbb{N}(0,1,1,0) + (1+6\,\mathbb{N})(0,1,1,0) + (2+0\,\mathbb{N})(0,0,1,0).$$

Furthermore, we have a factorization of $R^{\#}$ as given in Theorem 2.3: that is, $R^{\#} = R_0R_1R_2R_3$, where $R_0 = ((a,a,a,a)^5)^*$, $R_1 = ((\#,a,a,\#)^4)^*$, $R_2 = ((\#,a,a,\#)^6)^*(\#,a,a,\#)$, and $R_3 = (\#,\#,a,\#)^2$, with the same A_ℓ's as those defined above. The term $5(n,n,n,n)$ corresponds to R_0, the term $4(0,m,m,0)$ corresponds to R_1, the term $6(0,\ell,\ell,0) + (0,1,1,0)$ corresponds to R_2, and the term $(0,0,2,0)$ corresponds to R_3.

We need the following classical number-theoretic result (see [13], Thm. 1.0.1).

Theorem 2.9. *Let $a_1, \ldots, a_n$ be integers with $a_i \geq 2$ for $i = 1, \ldots, n$. If*

$$\gcd(a_1, \ldots, a_n) = 1,$$

then there exists a positive integer $F(a_1, \ldots, a_n)$ such that $F(a_1, \ldots, a_n)$ cannot be expressed as a non-negative linear combination of $a_1, \ldots, a_n$, but all integers greater than $F(a_1, \ldots, a_n)$ can be so expressed.

In the sequel we write $\mathbf{e}_i$ to denote the element of $\mathbb{N}^d$ that contains a 1 in its i-th component and 0's in all others.

Lemma 2.10. *A set $X \subseteq \mathbb{N}^d$ of the form (2.1) can be written as a union $A \cup B$, where A is made up of finite unions and intersections of sets having one of the forms (2.2)–(2.5) below and B is a finite intersection of sets of the form (2.2) or (2.3) below:*

$$\left\{ \sum_{\substack{i=1 \\ i \neq j}}^{d} n_i \mathbf{e}_i + (rn_j + s)\mathbf{e}_j : n_1, \ldots, n_d \in \mathbb{N}, n_j \geq N \right\} \tag{2.2}$$

where $1 \le j \le d$, *and* $r, s, N \in \mathbb{N}$;

$$\left\{ \sum_{\substack{i=1\\ i\neq j}}^{d} n_i \mathbf{e}_i + (n_k + rn_j + s)\mathbf{e}_j : n_1, \dots, n_d \in \mathbb{N}, n_j \ge N \right\} \tag{2.3}$$

where $1 \le j, k \le d$, $j \neq k$, *and* $r, s, N \in \mathbb{N}$;

$$\left\{ \sum_{\substack{i=1\\ i\neq j}}^{d} n_i \mathbf{e}_i + (rn_j + s)\mathbf{e}_j : n_1, \dots, n_d \in \mathbb{N}, n_j \in C \right\} \tag{2.4}$$

where $1 \le j \le d$, $r, s \in \mathbb{N}$, *and* $C \subseteq \mathbb{N}$ *is a finite set; or*

$$\left\{ \sum_{\substack{i=1\\ i\neq j}}^{d} n_i \mathbf{e}_i + (n_k + rn_j + s)\mathbf{e}_j : n_1, \dots, n_d \in \mathbb{N}, n_j \in C \right\} \tag{2.5}$$

where $1 \le j, k \le d$, $j \neq k$, *and* $r, s \in \mathbb{N}$, *and* $C \subseteq \mathbb{N}$ *is a finite set.*

Proof. Let X be a set of the form (2.1) where t, the A_ℓ's, the c_ℓ's, and the $b_{\ell,i}$'s are fixed and satisfy the conditions listed in Lemma 2.6. We will write $X = A \cup B$, where

$$B = \bigcap_{j=1}^{d} Y_j,$$

where each Y_j is either of the form (2.2) or (2.3), and A is made up of finite unions and intersections of sets of the forms (2.2)–(2.5).

First observe that if $j \in \{1, \dots, d\} \setminus A_0$ the set X contains only vectors whose j-th component is always 0. For each such j, we define

$$Y_j = \left\{ \sum_{\substack{i=1\\ i\neq j}}^{d} n_i \mathbf{e}_i + 0\mathbf{e}_j : n_1, \dots, n_d \in \mathbb{N} \right\},$$

which is of the form (2.2).

First consider the case where $A_0 = \ldots = A_t$. Define $j_1 < \ldots < j_{|A_0|}$ to be the elements of A_0. Define

$$Y_{j_1} = \left\{ \sum_{\substack{i=1\\ i\neq j_1}}^{d} n_i \mathbf{e}_i + (rn_{j_1} + s)\mathbf{e}_{j_1} : n_1, \dots, n_d \in \mathbb{N}, n_{j_1} \ge N \right\},$$

where $r = \gcd(c_0, \dots, c_t)$, $s = \sum_{\ell=0}^{t} b_{\ell,j_1}$, and $N - 1$ is the largest integer n such that rn cannot be written as a nonnegative integer linear combination of

$c_0, \ldots, c_t$ (note that by Theorem 2.9, N exists and is finite). Note that Y_{j_1} is of the form (2.2).

Define

$$Y'_{j_1} = \left\{ \sum_{\substack{i=1 \\ i \neq j_1}}^{d} n_i \mathbf{e}_i + (rn_{j_1} + s)\mathbf{e}_{j_1} : n_1, \ldots, n_d \in \mathbb{N}, n_{j_1} \in C \right\},$$

where C is the set of all nonnegative integers $n < N$ such that rn can be written as a nonnegative integer linear combination of $c_0, \ldots, c_t$. Note that Y'_{j_1} is of the form (2.4).

For $k \in \{2, \ldots, |A_0|\}$, define

$$Y_{j_k} = \left\{ \sum_{\substack{i=1 \\ i \neq j_k}}^{d} n_i \mathbf{e}_i + n_{j_{k-1}} \mathbf{e}_{j_k} : n_1, \ldots, n_d \in \mathbb{N} \right\},$$

which is of the form (2.3).

The set X can be written as the union $A \cup B$ where

$$B = \bigcap_{j \in \{1,\ldots,d\} \setminus A_0} Y_j \quad \cap \bigcap_{k \in \{1,\ldots,|A_0|\}} Y_{j_k}$$

and

$$A = \bigcap_{j \in \{1,\ldots,d\} \setminus A_0} Y_j \quad \cap \bigcap_{k \in \{2,\ldots,|A_0|\}} Y_{j_k} \quad \cap \quad Y'_{j_1}.$$

Now consider the case where there is at least one index ℓ such that $A_\ell \setminus A_{\ell+1} \neq \emptyset$. Define $\ell_1 < \ldots < \ell_{t'}$ to be the indices of the sets A_ℓ satisfying $A_{\ell_k} \setminus A_{\ell_k+1} \neq \emptyset$ for each $k \in \{1, \ldots, t'\}$. We clearly have $1 \leq t' \leq t$ and $0 \leq \ell_{t'} < t$.

Define $d_1 = |A_{\ell_1} \setminus A_{\ell_1+1}|$ and $j_{1,1} < \ldots < j_{1,d_1}$ to be the elements of $A_{\ell_1} \setminus A_{\ell_1+1}$. Define

$$Y_{j_{1,1}} = \left\{ \sum_{\substack{i=1 \\ i \neq j_{1,1}}}^{d} n_i \mathbf{e}_i + (r_1 n_{j_{1,1}} + s_1)\mathbf{e}_{j_{1,1}} : n_1, \ldots, n_d \in \mathbb{N}, n_{j_{1,1}} \geq N_1 \right\},$$

where $r_1 = \gcd(c_0, \ldots, c_{\ell_1})$, $s_1 = \sum_{\ell=0}^{\ell_1} b_{\ell, j_{1,1}}$, and $N_1 - 1$ is the largest integer n such that $r_1 n$ cannot be written as a nonnegative integer linear combination of $c_0, \ldots, c_{\ell_1}$. Note that $Y_{j_{1,1}}$ is of the form (2.2).

Define

$$Y'_{j_{1,1}} = \left\{ \sum_{\substack{i=1 \\ i \neq j_{1,1}}}^{d} n_i \mathbf{e}_i + (r_1 n_{j_{1,1}} + s_1)\mathbf{e}_{j_{1,1}} : n_1, \ldots, n_d \in \mathbb{N}, n_{j_{1,1}} \in C_1 \right\},$$

where C_1 is the set of all nonnegative integers $n < N_1$ such that $r_1 n$ can be written as a nonnegative integer linear combination of $c_0, \ldots, c_{\ell_1}$. Note that $Y'_{j_{1,1}}$ is of the form (2.4).

For $k \in \{2, \ldots, d_1\}$, define

$$Y_{j_{1,k}} = \left\{ \sum_{\substack{i=1\\ i \neq j_{1,k}}}^{d} n_i \mathbf{e}_i + n_{j_{1,k-1}} \mathbf{e}_{j_{1,k}} : n_1, \ldots, n_d \in \mathbb{N} \right\},$$

which is of the form (2.3).

Define $d_2 = |A_{\ell_2} \setminus A_{\ell_2+1}|$ and $j_{2,1} < \ldots < j_{2,d_2}$ to be the elements of $A_{\ell_2} \setminus A_{\ell_2+1}$. Define

$$Y_{j_{2,1}} = \left\{ \sum_{\substack{i=1\\ i \neq j_{2,1}}}^{d} n_i \mathbf{e}_i + (n_{j_{1,1}} + r_2 n_{j_{2,1}} + s_2) \mathbf{e}_{j_{2,1}} : n_1, \ldots, n_d \in \mathbb{N}, n_{j_{2,1}} \geq N_2 \right\},$$

where $r_2 = \gcd(c_{\ell_1+1}, \ldots, c_{\ell_2})$, $s_2 = \sum_{\ell=\ell_1+1}^{\ell_2} b_{\ell, j_{2,1}}$, and $N_2 - 1$ is the largest integer n such that $r_2 n$ cannot be written as a nonnegative integer linear combination of $c_{\ell_1+1}, \ldots, c_{\ell_2}$. Note that $Y_{j_{2,1}}$ is of the form (2.3).

Define

$$Y'_{j_{2,1}} = \left\{ \sum_{\substack{i=1\\ i \neq j_{2,1}}}^{d} n_i \mathbf{e}_i + (n_{j_{1,1}} + r_2 n_{j_{2,1}} + s_2) \mathbf{e}_{j_{2,1}} : n_1, \ldots, n_d \in \mathbb{N}, n_{j_{2,1}} \in C_2 \right\},$$

where C_2 is the set of all nonnegative integers $n < N_2$ such that $r_2 n$ can be written as a nonnegative integer linear combination of $c_{\ell_1+1}, \ldots, c_{\ell_2}$. Note that $Y'_{j_{2,1}}$ is of the form (2.5).

For $k \in \{2, \ldots, d_2\}$, define

$$Y_{j_{2,k}} = \left\{ \sum_{\substack{i=1\\ i \neq j_{2,k}}}^{d} n_i \mathbf{e}_i + n_{j_{2,k-1}} \mathbf{e}_{j_{2,k}} : n_1, \ldots, n_d \in \mathbb{N} \right\},$$

which is of the form (2.3).

We continue in this manner to define d_p, $Y_{j_{p,k}}$, and $Y'_{j_{p,1}}$ for all $p \in \{1, \ldots, t'\}$ and $k \in \{1, \ldots, d_p\}$. Finally observe that we have $A_{\ell_{t'}} \setminus A_{\ell_{t'}+1} \neq \emptyset$ and $A_{\ell_{t'}+1} = \ldots = A_t$. Define $d_{t'+1} = |A_t|$ and $j_{t'+1,1} < \ldots < j_{t'+1,d_{t'+1}}$ to be the elements of

A_t. Define

$$Y_{j_{t'+1,1}} = \left\{ \sum_{\substack{i=1\\ i\neq j_{t'+1,1}}}^{d} n_i\mathbf{e}_i + (n_{j_{t',1}} + r_{t'+1}n_{j_{t'+1,1}} + s_{t'+1})\mathbf{e}_{j_{t'+1,1}} : n_1,\dots,n_d \in \mathbb{N}, n_{j_{t'+1,1}} \geq N_{t'+1} \right\},$$

where $r_{t'+1} = \gcd(c_{\ell_{t'}+1},\dots,c_t)$, $s_{t'+1} = \sum_{\ell=\ell_{t'}+1}^{\ell_t} b_{\ell,j_{t'+1,1}}$, and $N_{t'+1}-1$ is the largest integer n such that $r_{t'+1}n$ cannot be written as a nonnegative integer linear combination of $c_{\ell_{t'}+1},\dots,c_t$. Again note that $Y_{j_{t'+1,1}}$ is of the form (2.3).

Define

$$Y'_{j_{t'+1,1}} = \left\{ \sum_{\substack{i=1\\ i\neq j_{t'+1,1}}}^{d} n_i\mathbf{e}_i + (n_{j_{t',1}} + r_{t'+1}n_{j_{t'+1,1}} + s_{t'+1})\mathbf{e}_{j_{t'+1,1}} : n_1,\dots,n_d \in \mathbb{N}, n_{j_{t'+1,1}} \in C_{t'+1} \right\},$$

where $C_{t'+1}$ is the set of all nonnegative integers $n < N_{t'+1}$ such that $r_{t'+1}n$ can be written as a nonnegative integer linear combination of $c_{\ell_{t'}+1},\dots,c_t$. Note that $Y'_{j_{t'+1,1}}$ is of the form (2.5).

For $k \in \{2,\dots,d_{t'+1}\}$, define

$$Y_{j_{t'+1,k}} = \left\{ \sum_{\substack{i=1\\ i\neq j_{t'+1,k}}}^{d} n_i\mathbf{e}_i + n_{j_{t'+1,k-1}}\mathbf{e}_{j_{t'+1,k}} : n_1,\dots,n_d \in \mathbb{N} \right\},$$

which is of the form (2.3).

The set X can be written as the union $A \cup B$ where

$$B = \bigcap_{j\in\{1,\dots,d\}\setminus A_0} Y_j \quad \cap \bigcap_{\substack{p\in\{1,\dots,t'+1\}\\ k\in\{1,\dots,d_p\}}} Y_{j_{p,k}}$$

and

$$A = \bigcap_{j\in\{1,\dots,d\}\setminus A_0} Y_j \quad \cap$$
$$\bigcup_{p\in\{1,\dots,t'+1\}} \left(Y'_{j_{p,1}} \quad \cap \bigcap_{\substack{q\in\{1,\dots,t'+1\}\setminus\{p\}\\ k\in\{1,\dots,d_q\}}} Y_{j_{q,k}} \quad \cap \bigcap_{k\in\{2,\dots,d_p\}} Y_{j_{p,k}} \right). \quad \square$$

Example 2.11. We continue Example 2.8. We will write $X = A \cup B$ as in Lemma 2.10. The A_ℓ's are not all the same, so we can define $t' = 2$, $\ell_1 = 0 < \ell_2 = 2$ as in the proof of Lemma 2.10.

We have $d_1 = |A_0 \setminus A_1| = 2$, $j_{1,1} = 1$ and $j_{1,2} = 4$. We also have $r_1 = \gcd(c_0) = \gcd(5) = 5$ and $s_1 = 0$, and hence $N_1 = 0$. Therefore,

$$Y_1 = \{n_2\mathbf{e_2} + n_3\mathbf{e_3} + n_4\mathbf{e_4} + (5n_1 + 0)\mathbf{e_1} : n_1, n_2, n_3, n_4 \in \mathbb{N}, n_1 \geq 0\},$$

$$Y'_1 = \{n_2\mathbf{e_2} + n_3\mathbf{e_3} + n_4\mathbf{e_4} + (5n_1 + 0)\mathbf{e_1} : n_2, n_3, n_4 \in \mathbb{N}, n_1 \in C_1\} = \emptyset,$$

since $C_1 = \emptyset$, and

$$Y_4 = \{n_1\mathbf{e_1} + n_2\mathbf{e_2} + n_3\mathbf{e_3} + n_1\mathbf{e_4} : n_1, n_2, n_3 \in \mathbb{N}\}.$$

Next we have $d_2 = |A_2 \setminus A_3| = 1$ and $j_{2,1} = 2$. We also have $r_2 = \gcd(c_1, c_2) = \gcd(4, 6) = 2$ and $s_2 = b_{1,2} + b_{2,2} = 0 + 1 = 1$, and hence $N_2 = 2$. Therefore,

$$Y_2 = \{n_1\mathbf{e_1} + n_3\mathbf{e_3} + n_4\mathbf{e_4} + (n_1 + 2n_2 + 1)\mathbf{e_2} : n_1, n_2, n_3 \in \mathbb{N}, n_2 \geq 2\},$$

and

$$\begin{aligned} Y'_2 &= \{n_1\mathbf{e_1} + n_3\mathbf{e_3} + n_4\mathbf{e_4} + (n_1 + 2n_2 + 1)\mathbf{e_2} : n_1, n_3 \in \mathbb{N}, n_2 \in C_2\} \\ &= \{n_1\mathbf{e_1} + n_3\mathbf{e_3} + n_4\mathbf{e_4} + (n_1 + 1)\mathbf{e_2} : n_1, n_3 \in \mathbb{N}\}, \end{aligned}$$

since $C_2 = \{0\}$.

Finally, we have $d_3 = |A_3| = 1$ and $j_{3,1} = 3$. We also have $r_3 = \gcd(c_3) = \gcd(0) = 0$ and $s_3 = b_{3,3} = 2$, and hence $N_3 = 0$. Therefore,

$$Y_3 = \{n_1\mathbf{e_1} + n_2\mathbf{e_2} + n_4\mathbf{e_4} + (n_2 + 0n_3 + 2)\mathbf{e_3} : n_1, n_2, n_3 \in \mathbb{N}, n_3 \geq 0\},$$

and

$$Y'_3 = \{n_1\mathbf{e_1} + n_2\mathbf{e_2} + n_4\mathbf{e_4} + (n_2 + 0n_3 + 2)\mathbf{e_3} : n_1, n_2 \in \mathbb{N}, n_3 \in C_3\} = \emptyset,$$

since $C_3 = \emptyset$.

Hence $A = Y_1 \cap Y'_2 \cap Y_3 \cap Y_4$ and $B = Y_1 \cap Y_2 \cap Y_3 \cap Y_4$.

Lemma 2.12. *Let $k \in \mathbb{N}$ and let S be an abstract numeration system. The set $X = \{(n, n+k) : n \in \mathbb{N}\}$ is S-recognizable.*

Proof. The proof follows easily from known results and so we only give a sketch of the proof. Let $R = \text{rep}_S(X)$. To show that X is S-recognizable we must show that $R^{\#}$ is a regular language. Consider first the set $Y = \{(\text{rep}_S(n), \text{rep}_S(n+1)) : n \in \mathbb{N}\}$. If we interpret Y as the function mapping $\text{rep}_S(n)$ to $\text{rep}_S(n+1)$, then Y is the so-called *successor function* (see [1] or [11] for more on the successor function). From [2], Proposition 3 (see also [9], Prop. 2.6.7), we have that Y is a *synchronous relation*. In [9] synchronous relations are defined in terms of letter-to-letter transducers, but this definition is equivalent to the fact that the language $Y^{\#}$ is accepted by a finite automaton. Moreover, from [8] (see also [9], Thm. 2.6.6), we have that the composition of synchronous relations is again a synchronous relation. Hence R, which is the k-fold composition of Y with itself, is a synchronous relation. We conclude that $R^{\#}$ is a regular language, as required. □

Lemma 2.13. *A set $X \subseteq \mathbb{N}^d$ having one of the forms (2.2)–(2.5) defined in Lemma 2.10 is S-recognizable for any abstract numeration system S.*

Proof. We will give the proof for the cases where X is either of the form (2.2) or (2.3) (the other two cases are similar).

Let $S = (L, \Sigma, <)$ be an abstract numeration system and let $\mathcal{T}$ be a finite automaton accepting L. Let $R = \text{rep}_S(X)$. We will show that $R^{\#}$ is regular. That is, we will define a (nondeterministic) finite automaton $\mathcal{M}$ that accepts $R^{\#}$. Let $(w_1, \ldots, w_d)^{\#}$ be an arbitrary input to the automaton $\mathcal{M}$.

Suppose that X is of the form (2.2). That is,

$$X = \left\{ \sum_{\substack{i=1 \\ i \neq j}}^{d} n_i \mathbf{e}_i + (rn_j + s)\mathbf{e}_j : n_1, \ldots, n_d \in \mathbb{N}, n_j \geq N \right\},$$

where $1 \leq j \leq d$, and $r, s, N \in \mathbb{N}$. Suppose first that $r = 0$. In this case, the automaton $\mathcal{M}$ simulates $\mathcal{T}$ on $w_1, \ldots, w_{j-1}, w_{j+1}, \ldots, w_d$. The automaton $\mathcal{M}$ accepts its input if and only if $\mathcal{T}$ accepts $w_1, \ldots, w_{j-1}, w_{j+1}, \ldots, w_d$ and $w_j = \text{rep}_S(s)$.

Now suppose that $r > 0$. By increasing the value of N, we may, without loss of generality, assume that $s < r$. By Theorem 1.3 the language $\{\text{rep}_S(rn_j + s) : n_j \geq N\}$ is regular. Let $\mathcal{T}'$ be an automaton accepting $\{\text{rep}_S(rn_j + s) : n_j \geq N\}$. As before, the automaton $\mathcal{M}$ simulates $\mathcal{T}$ on $w_1, \ldots, w_{j-1}, w_{j+1}, \ldots, w_d$, but now also simulates $\mathcal{T}'$ on w_j. The automaton $\mathcal{M}$ accepts its input if and only if $\mathcal{T}$ accepts $w_1, \ldots, w_{j-1}, w_{j+1}, \ldots, w_d$ and $\mathcal{T}'$ accepts w_j.

Next suppose that X is of the form (2.3). That is,

$$\left\{ \sum_{\substack{i=1 \\ i \neq j}}^{d} n_i \mathbf{e}_i + (n_k + rn_j + s)\mathbf{e}_j : n_1, \ldots, n_d \in \mathbb{N}, n_j \geq N \right\},$$

where $1 \le j, k \le d$, $j \ne k$, and $r, s, N \in \mathbb{N}$. Again, suppose first that $r = 0$. By Lemma 2.12, the language $\{(\mathrm{rep}_S(n_k), \mathrm{rep}_S(n_k+s))^\# : n_k \in \mathbb{N}\}$ is regular. Let $\mathcal{T}''$ be a finite automaton accepting this language. The automaton $\mathcal{M}$ simulates $\mathcal{T}$ on each of the words in $\{w_1, \ldots, w_d\} \setminus \{w_j, w_k\}$. Simultaneously, the automaton $\mathcal{M}$ simulates $\mathcal{T}''$ on the pair $(w_k, w_j)^\#$. The automaton $\mathcal{M}$ accepts its input if and only if $\mathcal{T}$ accepts $\{w_1, \ldots, w_d\} \setminus \{w_j, w_k\}$ and $\mathcal{T}''$ accepts $(w_k, w_j)^\#$.

Now suppose that $r > 0$. Again, without loss of generality, we may assume that $s < r$. Using the same ideas as in the proof of [12], Theorem 3.3.1, it is not hard to see that the language

$$\{(\mathrm{rep}_S(m), \mathrm{rep}_S(n))^\# : m, n \in \mathbb{N} \text{ and } (n - m) \equiv s \pmod r\}$$

is regular. Let $\mathcal{Z}$ be an automaton accepting this language. Let $\mathcal{Z}'$ be an automaton accepting the language $\{(\mathrm{rep}_S(n_k), \mathrm{rep}_S(n_k + rN + s))^\# : n_k \in \mathbb{N}\}$ (since $rN + s$ is a constant, we may apply Lem. 2.12).

The automaton $\mathcal{M}$ simulates $\mathcal{T}$ on each of the words in $\{w_1, \ldots, w_d\} \setminus \{w_j, w_k\}$. Simultaneously, the automaton $\mathcal{M}$ simulates $\mathcal{Z}$ on the pair $(w_k, w_j)^\#$.

The automaton $\mathcal{M}$ also nondeterministically "guesses" a word $v = b_1 \ldots b_{|v|}$ and simulates $\mathcal{Z}'$ on the pair $(w_k, v)^\#$. This "guess" works as follows. Let $w_k = a_1 \ldots a_{|w_k|}$, where each $a_i \in \Sigma$. For each $i = 1, \ldots, |w_k|$, we simulate $\mathcal{Z}'$ by nondeterministically choosing to follow one of the transitions of $\mathcal{Z}'$ labeled (a_i, b_i), where $b_i \in \Sigma$; and for $i > |w_k|$ (*i.e.*, w_k has been completely read), the simulation may make a nondeterministic choice among transitions of the form $(\#, b_i)$, where $b_i \in \Sigma$. This nondeterministic choice of b_i at each step of the simulation is what defines the "guessed" word v. Note that if $\mathcal{Z}'$ accepts $(w_k, v)^\#$, then $\mathrm{val}_s(v) = \mathrm{val}_S(w_k) + rN + s$. As this nondeterministic simulation is performed, the automaton $\mathcal{M}$ also simultaneously verifies that w_j is greater than or equal to (in the radix order) the guessed word v.

The automaton $\mathcal{M}$ accepts its input if and only if

- $\mathcal{T}$ accepts each of the words in $\{w_1, \ldots, w_d\} \setminus \{w_j, w_k\}$;
- $\mathcal{Z}$ accepts $(w_k, w_j)^\#$ (and hence $\mathrm{val}(w_j) - \mathrm{val}(w_k) \equiv s \pmod r$);
- $\mathcal{Z}'$ accepts $(w_k, v)^\#$ for some guessed word v as described above (and hence $\mathrm{val}_S(v) = \mathrm{val}_S(w_k) + rN + s$); and
- w_j is greater than or equal to v in the radix order (and hence $\mathrm{val}_S(w_j) \ge \mathrm{val}_S(v)$).

The last three of these conditions guarantee that $\mathrm{val}_S(w_j) = \mathrm{val}_S(w_k) + rn_j + s$ for some $n_j \ge N$.

This completes the proof for the cases where X is either of the form (2.2) or (2.3). As previously stated, we omit the details for the other two cases since they are similar. □

We are ready for the proof of Theorem 1.9.

Proof of Theorem 1.9. One direction is clear: if X is S-recognizable for all abstract numeration systems S, then it is certainly 1-recognizable.

To prove the other direction, suppose that X is 1-recognizable. The result now follows from Lemmas 2.6, 2.10, and 2.13. □

Acknowledgements. The work contained in the paper came about in response to a question posed by Jacques Sakarovitch during Michel Rigo's presentation of his thesis required for the "habilitation à diriger des recherches" in France. We thank Jacques Sakarovitch for his question and we thank Michel Rigo for presenting the problem to us. We also thank Christophe Reutenauer for the idea described in Remark 2.7.

References

[1] P.-Y. Angrand and J. Sakarovitch, Radix enumeration of rational languages. *RAIRO–Theor. Inf. Appl.* **44** (2010) 19–36.

[2] V. Berthé, C. Frougny, M. Rigo and J. Sakarovitch, On the cost and complexity of the successor function, in *Proceedings of WORDS 2007*. P. Arnoux, N. Bédaride and J. Cassaigne Eds., CIRM, Luminy, Marseille (2007).

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

[4] É. Charlier, T. Kärki and M. Rigo, Multidimensional generalized automatic sequences and shape-symmetric morphic words. *Discrete Math.* **310** (2010) 1238–1252.

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

[6] S. Eilenberg, *Automata, languages, and machines* **A**, *Pure and Applied Mathematics* **58**. Academic Press, New York (1974).

[7] S. Eilenberg, C.C. Elgot and J.C. Shepherdson, Sets recognised by n-tape automata. *J. Algebra* **13** (1969) 447–464.

[8] Ch. Frougny and J. Sakarovitch, Synchronized rational relations of finite and infinite words. *Theoret. Comput. Sci.* **108** (1993) 45–82.

[9] Ch. Frougny and J. Sakarovitch, Number representation and finite automata, in *Combinatorics, Automata, and Number Theory*, *Encyclopedia of Mathematics and its Applications* **135**. V. Berthé and M. Rigo Eds., Cambridge (2010).

[10] S. Ginsburg and E.H. Spanier, Semigroups, Presburger formulas and languages. *Pac. J. Math.* **16** (1966) 285–296.

[11] P. Lecomte and M. Rigo, Numeration systems on a regular language. *Theor. Comput. Syst.* **34** (2001) 27–44.

[12] P. Lecomte and M. Rigo, Abstract numeration systems, in *Combinatorics, Automata, and Number Theory*, *Encyclopedia of Mathematics and its Applications* **135**. V. Berthé and M. Rigo Eds., Cambridge (2010).

[13] J. Ramírez Alfonsín, *The Diophantine Frobenius problem*, *Oxford Lecture Series in Mathematics and its Applications* **30**. Oxford (2005).

[14] M. Rigo and A. Maes, More on generalized automatic sequences. *J. Autom. Lang. Comb.* **7** (2002) 351–376.

[15] S. Rubin, *Automatic Structures*. Ph.D. thesis, University of Auckland, New Zealand (2004).

[16] A.L. Semenov, The Presburger nature of predicates that are regular in two number systems. *Sibirsk. Math. Ž.* **18** (1977) 403–418, 479 (in Russian). English translation in *Sib. J. Math.* **18** (1977) 289–300.

Communicated by G. Richomme.
Received November 2, 2010. Accepted June 29, 2011.