

A CHARACTERIZATION OF MULTIDIMENSIONAL S -AUTOMATIC SEQUENCES

EMILIE CHARLIER, TOMI KÄRKI*, AND MICHEL RIGO

ABSTRACT. An infinite word is S -automatic if, for all $n \geq 0$, its $(n + 1)$ st letter is the output of a deterministic automaton fed with the representation of n in the considered numeration system S . In this extended abstract, we consider an analogous definition in a multidimensional setting and present the connection to the shape-symmetric infinite words introduced by Arnaud Maes. More precisely, for $d \geq 2$, we state that a multidimensional infinite word $x : \mathbb{N}^d \rightarrow \Sigma$ over a finite alphabet Σ is S -automatic for some abstract numeration system S built on a regular language containing the empty word if and only if x is the image by a coding of a shape-symmetric infinite word.

1. INTRODUCTION

Let $k \geq 2$. An infinite word $x = (x_n)_{n \geq 0}$ is k -automatic if for all $n \geq 0$, x_n is obtained by feeding a deterministic finite automaton with output (DFAO for short) with the k -ary representation of n . The theorem of Cobham from 1972 [4] states that a word $(x_n)_{n \geq 0}$ is k -automatic if and only if it is generated by a morphism of constant length k and possibly a final application of a coding.

An extension of this result obtained by Salon [12] deals with a multidimensional setting and an integer base k numeration system. In this framework, images of letters by a morphism are hypercubes of constant size k . Another possible extension of Cobham's theorem in the unidimensional setting is to relax the hypothesis on the constant length of the morphism: a word $(x_n)_{n \geq 0}$ is generated by an arbitrary prolongable morphism and possibly a final coding if and only if it is S -automatic [10, 11]. The $(n + 1)$ st letter of an S -automatic word is obtained by feeding a DFAO with the representation of n in the considered abstract numeration system S [5].

Here this latter result is extended to a multidimensional setting. The notion of an S -automatic multidimensional word is easy to define. However, difficulties occur when one has to iterate a morphism where the images of the letters are not anymore hypercubes of constant size.

A longer version of this note containing detailed proofs has been submitted for publication [3].

The next sections are devoted to the presentation of the different concepts occurring in this note. Then we state our main result and we conclude this note by some perspectives.

2. ABSTRACT NUMERATION SYSTEMS

For a finite alphabet Σ , Σ^* denotes the free monoid generated by Σ having concatenation of words as product and the empty word ε as neutral element. For a word $w = w_0 \cdots w_{\ell-1}$, $\ell \geq 0$, and $w_j \in \Sigma$, $|w|$ denotes its length ℓ . Let $(\Sigma, <)$ be a totally ordered alphabet and u, v be two words over Σ . We say that u is *genealogically less* than v , and we write $u \prec v$ if either $|u| < |v|$ (i.e., u is of shorter length than v) or $|u| = |v|$ and there exist $p, s, t \in \Sigma^*$, $a, b \in \Sigma$ such that $u = pas$,

* Supported by Osk. Huttunen Foundation.

$v = pbt$ and $a < b$ (i.e., u is lexicographically less than v). Let us also mention that we have taken the convention that all finite and infinite words and pictures have indices starting from 0.

Definition 1. An *abstract numeration system* [5] is a triple $S = (L, \Sigma, <)$ where L is an infinite regular language over a totally ordered finite alphabet $(\Sigma, <)$. Enumerating the words of L using the genealogical ordering $<$ induced by the ordering $<$ of Σ gives a one-to-one correspondence $\text{rep}_S : \mathbb{N} \rightarrow L$ mapping any non-negative integer n onto the $(n + 1)$ st word in L .

Any regular language is accepted by a deterministic finite automaton defined as follows. A *deterministic finite automaton* \mathcal{A} (DFA for short) is given by $\mathcal{A} = (Q, q_0, \Sigma, \delta, F)$ where Q is the finite *set of states*, $q_0 \in Q$ is the *initial state*, $\delta : Q \times \Sigma \rightarrow Q$ is the *transition function* and $F \subseteq Q$ is the *set of final states*. The function δ can be extended to $Q \times \Sigma^*$ by $\delta(q, \varepsilon) = q$ for all $q \in Q$ and $\delta(q, aw) = \delta(\delta(q, a), w)$ for all $q \in Q$, $a \in \Sigma$ and $w \in \Sigma^*$. A word $w \in \Sigma^*$ is *accepted by* \mathcal{A} if $\delta(q_0, w) \in F$. The *language accepted by* \mathcal{A} is the set of the accepted words. A *deterministic finite automaton with output* (DFAO for short) $\mathcal{B} = (Q, q_0, \Sigma, \delta, \Gamma, \tau)$ is a generalization of a DFA where Γ is the *output alphabet* and $\tau : Q \rightarrow \Gamma$ is the *output function*. The *output corresponding to the input* $w \in \Sigma^*$ is $\tau(\delta(q_0, w))$.

3. S -AUTOMATIC MULTIDIMENSIONAL INFINITE WORDS

Let $d \geq 1$. To work with d -tuples of words of the same length, we introduce the following map.

Definition 2. If w_1, \dots, w_d are finite words over the alphabet Σ , the map $(\cdot)^\# : (\Sigma^*)^d \rightarrow ((\Sigma \cup \{\#\})^d)^*$ is defined as

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

where $m = \max\{|w_1|, \dots, |w_d|\}$.

Definition 3. A *d -dimensional infinite word* over the alphabet Γ is a map $x : \mathbb{N}^d \rightarrow \Gamma$. We use notation x_{n_1, \dots, n_d} or $x(n_1, \dots, n_d)$ to denote the value of x at (n_1, \dots, n_d) . Such a word is said to be *S -automatic* if there exist an abstract numeration system $S = (L, \Sigma, <)$ and a deterministic finite automaton with output $\mathcal{A} = (Q, q_0, (\Sigma \cup \{\#\})^d, \delta, \Gamma, \tau)$ such that, for all $n_1, \dots, n_d \geq 0$,

$$\tau(\delta(q_0, (\text{rep}_S(n_1), \dots, \text{rep}_S(n_d))^\#)) = x_{n_1, \dots, n_d}.$$

This notion was introduced in [11] (see also [9]) as a natural generalization of the multidimensional k -automatic sequences introduced in [12].

Example 4. Consider the abstract numeration system

$$S = (\{a, ba\}^*\{\varepsilon, b\}, \{a, b\}, a < b)$$

and the DFAO depicted in Figure 1. Since this automaton is fed with entries of the form $(\text{rep}_S(n_1), \text{rep}_S(n_2))^\#$, we do not consider the transitions of label $(\#, \#)$. If the outputs of the DFAO are considered to be the states themselves, then we produce the bidimensional infinite S -automatic word given in Figure 2.

4. MULTIDIMENSIONAL MORPHISM

This section is dedicated to present the notions of a multidimensional morphism and shape-symmetry as they were introduced by A. Maes mainly in connection with the decidability question of logical theories [6, 7, 8].

If $i \leq j$ are integers, $\llbracket i, j \rrbracket$ denotes the interval of integers $\{i, i + 1, \dots, j\}$. Let $d \geq 1$. If $\mathbf{n} \in \mathbb{N}^d$ and $i \in \{1, \dots, d\}$, then n_i is the i th component of \mathbf{n} . Let \mathbf{m}

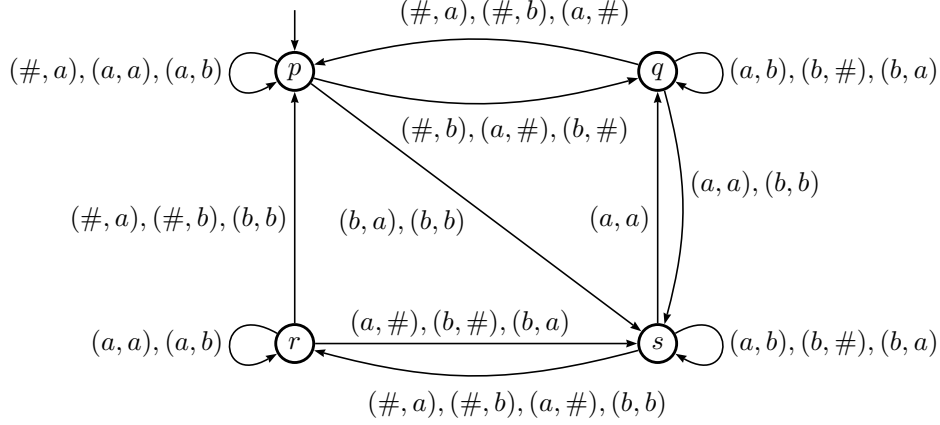


FIGURE 1. A deterministic finite automaton with output.

	ω	a	b	aa	ab	ba	aaa	aab	\dots
ε	p	q	q	p	q	p	q	q	\dots
a	p	p	s	s	q	s	p	s	
b	q	p	s	q	s	q	p	s	
aa	p	p	s	p	s	q	q	s	
ab	q	p	s	p	s	s	s	r	
ba	p	s	q	p	s	q	s	q	
aaa	p	p	s	p	s	q	p	s	
aab	q	p	s	p	s	s	p	s	
\vdots	\vdots								\ddots

FIGURE 2. A bidimensional infinite S -automatic word.

and \mathbf{n} be two d -tuples in \mathbb{N}^d . We write $\mathbf{m} \leq \mathbf{n}$ (resp. $\mathbf{m} < \mathbf{n}$), if $m_i \leq n_i$ (resp. $m_i < n_i$) for all $i = 1, \dots, d$. For $\mathbf{n} \in \mathbb{N}^d$ and $j \in \mathbb{N}$, $\mathbf{n} + j := (n_1 + j, \dots, n_d + j)$. In particular, we set $\mathbf{0} := (0, \dots, 0)$ and $\mathbf{1} := (1, \dots, 1)$. If $j \cdot \mathbf{1} \leq \mathbf{n}$, then we set $\mathbf{n} - j := (n_1 - j, \dots, n_d - j)$.

Definition 5. Let s_1, \dots, s_d be positive integers or ∞ . A d -dimensional picture over the alphabet Σ is a map x with domain $\llbracket 0, s_1 - 1 \rrbracket \times \dots \times \llbracket 0, s_d - 1 \rrbracket$ taking values in Σ . By convention, if $s_i = \infty$ for some i , then $\llbracket 0, s_i - 1 \rrbracket = \mathbb{N}$. If x is such a picture, we write $|x|$ for the d -tuple $(s_1, \dots, s_d) \in (\mathbb{N} \cup \{\infty\})^d$ which is called the *shape* of x . We denote by ε_d the d -dimensional picture of shape $(0, \dots, 0)$. Note that $\varepsilon_1 = \varepsilon$. If $\mathbf{n} = (n_1, \dots, n_d)$ belongs to the domain of x , we indifferently use the notation x_{n_1, \dots, n_d} , $x_{\mathbf{n}}$, $x(n_1, \dots, n_d)$ or $x(\mathbf{n})$. Let x be a d -dimensional picture. If for all $i \in \{1, \dots, d\}$, $|x|_i < \infty$, then x is said to be *bounded*. The set of d -dimensional bounded pictures over Σ is denoted by $B_d(\Sigma)$. A bounded picture x is a *square* of size $c \in \mathbb{N}$ if $|x| = c \cdot \mathbf{1}$.

Definition 6. Let x be a d -dimensional picture. If $\mathbf{0} \leq \mathbf{s} \leq \mathbf{t} \leq |x| - \mathbf{1}$, then $x[\mathbf{s}, \mathbf{t}]$ is said to be a *factor* of x and is defined as the picture y of shape $\mathbf{t} - \mathbf{s} + \mathbf{1}$ given by $y(\mathbf{n}) = x(\mathbf{n} + \mathbf{s})$ for all $\mathbf{n} \in \mathbb{N}^d$ such that $\mathbf{n} \leq \mathbf{t} - \mathbf{s}$. For any $\mathbf{u} \in \mathbb{N}^d$, the set of factors of x of shape \mathbf{u} is denoted by $\text{Fact}_{\mathbf{u}}(x)$.

Example 7. Consider the bidimensional (bounded) picture of shape $(5, 2)$,

$$x = \begin{array}{|c|c|c|c|c|} \hline a & b & a & a & b \\ \hline c & d & b & c & d \\ \hline \end{array}.$$

We have

$$x[(0, 0), (1, 1)] = \begin{array}{|c|c|} \hline a & b \\ \hline c & d \\ \hline \end{array} \quad \text{and} \quad x[(2, 0), (4, 1)] = \begin{array}{|c|c|c|} \hline a & a & b \\ \hline b & c & d \\ \hline \end{array}.$$

For instance, $\text{Fact}_1(x) = \{a, b, c, d\}$ and

$$\text{Fact}_{(3,2)}(x) = \left\{ \begin{array}{|c|c|c|} \hline a & b & a \\ \hline c & d & b \\ \hline \end{array}, \begin{array}{|c|c|c|} \hline b & a & a \\ \hline d & b & c \\ \hline \end{array}, \begin{array}{|c|c|c|} \hline a & a & b \\ \hline b & c & d \\ \hline \end{array} \right\}.$$

Definition 8. Let x be a d -dimensional picture of shape $\mathbf{s} = (s_1, \dots, s_d)$. For all $i \in \{1, \dots, d\}$ and $k < s_i$, $x|_{i,k}$ is the $(d-1)$ -dimensional picture of shape

$$|x|_{\hat{i}} = \mathbf{s}_{\hat{i}} := (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_d)$$

defined by

$$x|_{i,k}(n_1, \dots, n_{i-1}, n_{i+1}, \dots, n_d) = x(n_1, \dots, n_{i-1}, k, n_{i+1}, \dots, n_d)$$

for all $0 \leq n_j < s_j$, $j \in \{1, \dots, d\} \setminus \{i\}$.

Definition 9. Let x, y be two d -dimensional pictures. If for some $i \in \{1, \dots, d\}$, $|x|_{\hat{i}} = |y|_{\hat{i}} = (s_1, \dots, s_{j-1}, s_{j+1}, \dots, s_d)$, then we define the *concatenation* of x and y in the direction i as the d -dimensional picture $x \odot^i y$ of shape $(s_1, \dots, s_{j-1}, |x|_i + |y|_i, s_{j+1}, \dots, s_d)$ satisfying

- (i) $x = (x \odot^i y)[\mathbf{0}, |x| - 1]$
- (ii) $y = (x \odot^i y)[(0, \dots, 0, |x|_i, 0, \dots, 0), (0, \dots, 0, |x|_i, 0, \dots, 0) + |y| - 1]$.

The d -dimensional empty word ε_d is a word of shape $\mathbf{0}$. We extend the definition to the concatenation of ε_d and any d -dimensional word x in the direction $i \in \{1, \dots, d\}$ by

$$\varepsilon_d \odot^i x = x \odot^i \varepsilon_d = x.$$

Especially, $\varepsilon_d \odot^i \varepsilon_d = \varepsilon_d$.

Example 10. Consider the two bidimensional pictures

$$x = \begin{array}{|c|c|} \hline a & b \\ \hline c & d \\ \hline \end{array} \quad \text{and} \quad y = \begin{array}{|c|c|c|} \hline a & a & b \\ \hline b & c & d \\ \hline \end{array}$$

with shapes $|x| = (2, 2)$ and $|y| = (3, 2)$. Since $|x|_{\hat{1}} = |y|_{\hat{1}} = 2$, we get

$$x \odot^1 y = \begin{array}{|c|c|c|c|} \hline a & b & a & a & b \\ \hline c & d & b & c & d \\ \hline \end{array}.$$

But notice that $x \odot^2 y$ is not defined because $2 = |x|_{\hat{2}} \neq |y|_{\hat{2}} = 3$.

Let x be a d -dimensional picture and $\mu : \Sigma \rightarrow B_d(\Sigma)$ be a map. Note that μ cannot necessarily be extended to a morphism on Σ^* . Indeed, if x is a picture over Σ , $\mu(x)$ is not always well defined. Depending on the shapes of the images by μ of the letters in Σ , when trying to build $\mu(x)$ by concatenating the images $\mu(x_i)$ we can obtain “holes” or “overlaps”. Therefore, we introduce some restrictions on μ .

Definition 11. Let $\mu : \Sigma \rightarrow B_d(\Sigma)$ be a map and x be a d -dimensional picture such that

$$\forall i \in \{1, \dots, d\}, \forall k < |x|_i, \forall a, b \in \text{Fact}_1(x|_{i,k}) : |\mu(a)|_i = |\mu(b)|_i.$$

Then $\mu(x)$ is defined as

$$\mu(x) = \odot_{0 \leq n_1 < |x|_1}^1 \left(\dots \left(\odot_{0 \leq n_d < |x|_d}^d \mu(x(n_1, \dots, n_d)) \right) \right).$$

Note that the ordering of the products in the different directions is unimportant.

Example 12. Consider the map μ given by

$$a \mapsto \begin{bmatrix} a & a \\ b & d \end{bmatrix}, \quad b \mapsto \begin{bmatrix} c \\ b \end{bmatrix}, \quad c \mapsto \boxed{a \ a}, \quad d \mapsto \boxed{d}.$$

Let

$$x = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Since $|\mu(a)|_2 = |\mu(b)|_2 = 2$, $|\mu(c)|_2 = |\mu(d)|_2 = 1$, $|\mu(a)|_1 = |\mu(c)|_1 = 2$ and $|\mu(b)|_1 = |\mu(d)|_1 = 1$, $\mu(x)$ is well defined and given by

$$\mu(x) = \begin{bmatrix} a & a & c \\ b & d & b \\ a & a & d \end{bmatrix}.$$

But one can notice that $\mu^2(x)$ is not well defined.

Definition 13. Let $\mu : \Sigma \rightarrow B_d(\Sigma)$ be a map. If for all $a \in \Sigma$ and all $n \geq 0$, $\mu^n(a)$ is well defined from $\mu^{n-1}(a)$, then μ is said to be a d -dimensional morphism.

The usual notion of a *prolongable morphism* can be given in this multidimensional setting.

Definition 14. Let μ be a d -dimensional morphism. We say that μ is *prolongable on a* if a is a letter such that $(\mu(a))_{\mathbf{0}} = a$. Then the limit

$$w = \mu^\omega(a) := \lim_{n \rightarrow +\infty} \mu^n(a)$$

is well defined and $w = \mu(w)$ is a *fixed point* of μ . A d -dimensional infinite word x over Σ is said to be *purely morphic* if it is a fixed point of a d -dimensional morphism. It is said to be *morphic* if there exists a coding $\nu : \Gamma \rightarrow \Sigma$ (i.e., a letter-to-letter morphism) such that $x = \nu(y)$ for some purely morphic word y over Γ .

5. SHAPE-SYMMETRIC MORPHIC WORDS

The property of shape-symmetry is a natural generalization of uniform morphisms where all images are squares of the same shape [12].

Definition 15. Let $\mu : \Sigma \rightarrow B_d(\Sigma)$ be a d -dimensional morphism having the d -dimensional infinite word x as a fixed point. If for any permutation f of $\{1, \dots, d\}$ and for all $n_1, \dots, n_d > 0$, $|\mu(x(n_{f(1)}, \dots, n_{f(d)}))| = (s_{f(1)}, \dots, s_{f(d)})$ whenever $|\mu(x(n_1, \dots, n_d))| = (s_1, \dots, s_d)$, then x is said to be *shape-symmetric* (with respect to μ).

Remark 16. A. Maes showed that determining whether or not a map $\mu : \Sigma \rightarrow B_d(\Sigma)$ is a d -dimensional morphism is a decidable problem. Moreover he showed that if μ is prolongable on a letter a , then it is decidable whether or not the fixed point $\mu^\omega(a)$ is shape-symmetric [6, 7, 8].

Example 17. One can show that the following morphism has a fixed point $\mu^\omega(a)$ which is shape-symmetric.

$$\begin{aligned} \mu(a) = \mu(f) &= \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \mu(b) = \begin{bmatrix} e \\ c \end{bmatrix}, \quad \mu(c) = \boxed{e \ b}, \quad \mu(d) = \boxed{f}, \quad \mu(e) = \begin{bmatrix} e & b \\ g & d \end{bmatrix}, \\ \mu(g) &= \boxed{h \ b}, \quad \mu(h) = \begin{bmatrix} h & b \\ c & d \end{bmatrix}. \end{aligned}$$

We have represented in Figure 3 the beginning of the picture.

a	b	e	e	b	e	e	b	\dots	
c	d	c	g	d	g	c	g	d	
e	b	f	e	b	h	b	f	h	b
e	b	e	a	b	e	b	e	h	b
g	d	c	c	d	g	d	c	c	d
e	b	e	e	b	a	b	e	e	b
g	d	c	g	d	c	d	c	g	d
h	b	f	e	b	e	b	f	h	b
e	b	e	e	b	e	b	e	a	b
g	d	c	g	d	g	d	c	c	d
\vdots								\ddots	

FIGURE 3. A fixed point of μ .

6. MAIN RESULT

Using the above notions, we are now able to state our main result. A complete proof can be found in [3].

Theorem 18. *Let $d \geq 1$. The d -dimensional infinite word x is S -automatic for some abstract numeration system $S = (L, \Sigma, <)$ where $\varepsilon \in L$ if and only if x is the image by a coding of a shape-symmetric d -dimensional infinite word.*

7. PERSPECTIVES

In the context of this work, some natural questions arise. We point out some of them. For the sake of clarity, we write statements in the bidimensional case but each of them can be more generally stated in the d -dimensional case for $d \geq 2$.

- In the unidimensional case, S -automatic words are characterized in terms of finite S -kernels [11]. Does this characterization hold in the multidimensional setting?
- Multidimensional k -automatic words have been characterized in logical terms, see for instance [2]. Can such a characterization be extended to S -automatic words?
- Any bidimensional word $a = (a_{m,n})_{m,n \geq 0}$ over an alphabet Σ which is embedded into a finite field K can be associated with a formal power series $F(a) = \sum_{m,n \geq 0} a_{m,n} X^m Y^n$ over $K[[X, Y]]$. If a is S -automatic, can we derive some algebraic properties of $F(a)$, and conversely?
- As a particular case of the previous question, can something be said about the diagonal $D(a) = \sum_{m \geq 0} a_{m,m} X^m \in K[[X]]$?
- If S and T are two abstract numeration systems, one can easily adapt Definition 3 to define the notion of (S, T) -automatic word, i.e., $x_{m,n} = \tau(\delta(q_0, (\text{rep}_S(m), \text{rep}_T(n))^\#))$. Can these (S, T) -automatic words be characterized using morphisms?
- Peano enumeration of \mathbb{N}^2 is defined by $P: \mathbb{N}^2 \rightarrow \mathbb{N}: (m, n) \mapsto \frac{1}{2}(m+n)(m+n+1) + n$. Hence, to any bidimensional word $(a_{m,n})_{m,n \geq 0}$ corresponds a unidimensional word $(b_\ell)_{\ell \geq 0}$ where $b_\ell = a_{m,n}$ if $P(m, n) = \ell$. Suppose that $(a_{m,n})_{m,n \geq 0}$ is S -automatic. Does this imply that $(b_\ell)_{\ell \geq 0}$ is S -automatic too, and conversely? One could also consider other primitive recursive enumerations of \mathbb{N}^2 .

REFERENCES

- [1] J.-P. Allouche and J. Shallit, *Automatic Sequences: Theory, Applications, Generalizations*, Cambridge University Press, (2003).
- [2] V. Bruyère, G. Hansel, C. Michaux, and R. Villemaire, Logic and p -recognizable sets of integers, *Bull. Belg. Math. Soc.* **1** (1994), 191–238.
- [3] E. Charlier, T. Kärki, and M. Rigo, Multidimensional Generalized Automatic Sequences and Shape-Symmetric Morphic Words, submitted for publication.
- [4] A. Cobham, Uniform Tag Sequences, *Math. Systems Theory* **6** (1972), 164–192.
- [5] P.B.A. Lecomte, M. Rigo, Numeration Systems on a Regular Language, *Theory Comput. Syst.* **34** (2001), 27–44.
- [6] A. Maes, Decidability of the First-Order Theory of $\langle \mathbb{N}; <, P \rangle$ for Morphic Predicates P , Preprint 9806, Inst. für Informatik und Praktische Math., Christian-Albrechts-Univ. Kiel (1998).
- [7] A. Maes, An Automata-Theoretic Decidability Proof for First-Order Theory of $\langle \mathbb{N}, <, P \rangle$ with Morphic Predicate P , *J. Autom. Lang. Comb.* **4** (1999), 229–245.
- [8] A. Maes, Morphic Predicates and Applications to the Decidability of Arithmetic Theories, Ph.D. Thesis, Univ. Mons-Hainaut, (1999).
- [9] S. Nicolay, M. Rigo, About the Frequency of Letters in Generalized Automatic Sequences, *Theoret. Comp. Sci.* **374** (2007), 25–40.
- [10] M. Rigo, Generalization of Automatic Sequences for Numeration Systems on a Regular Language, *Theoret. Comp. Sci.* **244** (2000), 271–281.
- [11] M. Rigo and A. Maes, More on Generalized Automatic Sequences, *J. Autom., Lang. and Comb.* **7** (2002), 351–376.
- [12] O. Salon, Suites automatiques à multi-indices, *Séminaire de théorie des nombres de Bordeaux*, Exp. **4** (1986-1987), 4.01–4.27; followed by an Appendix by J. Shallit, 4-29A–4-36A.

INSTITUTE OF MATHEMATICS, UNIVERSITY OF LIÈGE, GRANDE TRAVERSE 12 (B 37), B-4000
LIÈGE, BELGIUM

E-mail address: {echarlier,T.Karki,M.Rigo}@ulg.ac.be