Abstract Numeration Systems

Émilie Charlier

Université libre de Bruxelles

Colloquium
Turku, May 7, 2012

Positional numeration systems

A positional numeration system (PNS) is given by a sequence of integers $U = (U_i)_{i\geq 0}$ such that

- $U_0 = 1$
- $\forall i\ U_i < U_{i+1}$
- $(U_{i+1}/U_i)_{i\geq 0}$ is bounded $\quad\rightarrow C_U = \sup_{i\geq 0}\lceil U_{i+1}/U_i \rceil$

The greedy U-representation of a positive integer n is the unique word $\text{rep}_U(n) = c_{\ell-1}\cdots c_0$ over $\Sigma_U = \{0,\ldots,C_U - 1\}$ satisfying

$$n = \sum_{i=0}^{\ell-1} c_i\, U_i,\ c_{\ell-1} \neq 0 \text{ and } \forall t\ \sum_{i=0}^{t} c_i U_i < U_{t+1}.$$

Recognizable sets of integers

A set $X \subseteq \mathbb{N}$ is U-recognizable or U-automatic if the subset $\mathrm{rep}_U(X) = \{\mathrm{rep}_U(x) \colon x \in X\}$ of Σ_U^* is regular.

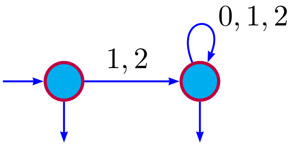
Integer base $b \geq 2$

$U = (b^i)_{i \geq 0}$

$\Sigma_b = \{0, \cdots, b-1\}$

$\mathcal{L}_b = \mathrm{rep}_b(\mathbb{N}) = \Sigma_b^* \setminus 0\,\Sigma_b^*$

$\mathrm{rep}_U,\ \Sigma_U \ \rightarrow\ \mathrm{rep}_b,\ \Sigma_b$



$\mathrm{rep}_3(\mathbb{N})$

$\mathbb{N}$ is 3-recognizable

27	9	3	1	
			ε	0
			1	1
			2	2
		1	0	3
		1	1	4
		1	2	5
		2	0	6
		2	1	7
		2	2	8
	1	0	0	9

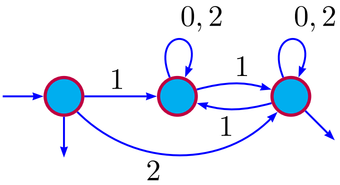
Integer base $b \geq 2$

$U = (b^i)_{i \geq 0}$

$\mathrm{rep}_U,\ \Sigma_U \ \rightarrow\ \mathrm{rep}_b,\ \Sigma_b$

$\Sigma_b = \{0, \cdots, b-1\}$

$\mathcal{L}_b = \mathrm{rep}_b(\mathbb{N}) = \Sigma_b^* \setminus 0\,\Sigma_b^*$



$\mathrm{rep}_{U_3}(2\,\mathbb{N})$

$2\,\mathbb{N}$ is 3-recognizable

27	9	3	1	
			ε	0
			1	1
			2	2
		1	0	3
		1	1	4
		1	2	5
		2	0	6
		2	1	7
		2	2	8
	1	0	0	9

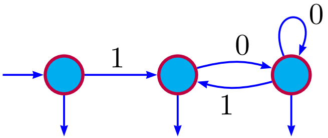
Fibonacci (or Zeckendorf) numeration system

Let $F = (F_i)_{i \geq 0} = (1, 2, 3, 5, 8, 13, 21, \ldots)$ be defined by

$$F_0 = 1,\ F_1 = 2 \ \text{ and } \ \forall i \in \mathbb{N},\ F_{i+2} = F_{i+1} + F_i.$$

$\Sigma_F = \{0, 1\}$

The factor 11 is forbidden :



$\mathrm{rep}_F(\mathbb{N}) = 1\{0, 01\}^* \cup \{\varepsilon\}$

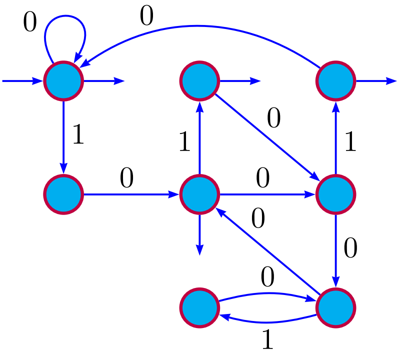
$\mathbb{N}$ is F-recognizable

13	8	5	3	2	1	
					ε	0
					1	1
				1	0	2
			1	0	0	3
			1	0	1	4
		1	0	0	0	5
		1	0	0	1	6
		1	0	1	0	7
	1	0	0	0	0	8

Fibonacci (or Zeckendorf) numeration system

Let $F = (F_i)_{i \geq 0} = (1, 2, 3, 5, 8, 13, 21, \dots)$ be defined by

$$F_0 = 1, \ F_1 = 2 \ \text{ and } \ \forall i \in \mathbb{N}, \ F_{i+2} = F_{i+1} + F_i.$$



$\mathrm{rep}_F(2\,\mathbb{N})$

$2\,\mathbb{N}$ is F-recognizable

13	8	5	3	2	1	
					ε	0
					1	1
				1	0	2
			1	0	0	3
			1	0	1	4
		1	0	0	0	5
		1	0	0	1	6
		1	0	1	0	7
	1	0	0	0	0	8

U-recognizability of $\mathbb{N}$

Is the set $\mathbb{N}$ U-recognizable? Otherwise stated, is the numeration language $\mathrm{rep}_U(\mathbb{N})$ regular? Not necessarily:

Theorem (Shallit 1994)

Let U be a PNS. If $\mathbb{N}$ is U-recognizable, then U is linear, i.e., it satisfies a linear recurrence relation over $\mathbb{Z}$.

Loraud (1995) and Hollander (1998) gave sufficient conditions for the numeration language to be regular : "The characteristic polynomial of the recurrence relation has a particular form".

U-recognizability of arithmetic progressions

Proposition

Let $U = (U_i)_{i\geq 0}$ *be a linear numeration system and let* $p, q \in \mathbb{N}$*.*

If $\mathbb{N}$ *is* U*-recognizable, then* $p + \mathbb{N}\, q$ *is* U*-recognizable and a DFA accepting* $\mathrm{rep}_U(p + \mathbb{N}\, q)$ *can be obtained efficiently.*

Consequently, any ultimately periodic set is U*-recognizable.*

Abstract numeration systems

An abstract numeration system (ANS) is a triple $S = (L, \Sigma, <)$ where L is an infinite regular language over a totally ordered alphabet $(\Sigma, <)$.
By enumerating the words of L w.r.t. the radix order $<_{rad}$ induced by $<$, we define a bijection :

$$\mathrm{rep}_S : \mathbb{N} \to L \qquad \mathrm{val}_S = \mathrm{rep}_S^{-1} : L \to \mathbb{N}.$$

A set $X \subseteq \mathbb{N}$ is S-recognizable if $\mathrm{rep}_S(X)$ is regular.

$L = \{a, b\}^* \quad \Sigma = \{a, b\} \quad a < b$

n	0	1	2	3	4	5	6	7	$\cdots$
rep(n)	ε	a	b	aa	ab	ba	bb	aaa	$\cdots$

$L = a^*b^* \quad \Sigma = \{a, b\} \quad a < b$

n	0	1	2	3	4	5	6	$\cdots$
rep(n)	ε	a	b	aa	ab	bb	aaa	$\cdots$

A generalization

ANS generalize PNS having a regular numeration language:

Let U be a PNS and let $x, y \in \mathbb{N}$. We have

$$x < y \Leftrightarrow \mathrm{rep}_U(x) <_{rad} \mathrm{rep}_U(y).$$

Example (Fibonacci)

$\mathrm{rep}_F(\mathbb{N}) = 1\{0, 01\}^* \cup \{\varepsilon\}$

$6 < 7$ and $1001 <_{rad} 1010$ (same length)

$6 < 8$ and $1001 <_{rad} 10000$ (different lengths)

13	8	5	3	2	1	
					ε	0
					1	1
				1	0	2
			1	0	0	3
			1	0	1	4
		1	0	0	0	5
		1	0	0	1	6
		1	0	1	0	7
	1	0	0	0	0	8

A generalization

ANS generalize PNS having a regular numeration language:

Let U be a PNS and let $x, y \in \mathbb{N}$. We have

$$x < y \Leftrightarrow \text{rep}_U(x) <_{rad} \text{rep}_U(y).$$

Example (Fibonacci)

$\text{rep}_F(\mathbb{N}) = 1\{0, 01\}^* \cup \{\varepsilon\}$

$6 < 7$ and $1001 <_{rad} 1010$
(same length)

$6 < 8$ and $1001 <_{rad} 10000$
(different lengths)

13	8	5	3	2	1	
					ε	0
					1	1
				1	0	2
			1	0	0	3
			1	0	1	4
		1	0	0	0	5
		1	0	0	1	6
		1	0	1	0	7
	1	0	0	0	0	8

A generalization

ANS generalize PNS having a regular numeration language:

Let U be a PNS and let $x, y \in \mathbb{N}$. We have

$$x < y \Leftrightarrow \mathrm{rep}_U(x) <_{rad} \mathrm{rep}_U(y).$$

Example (Fibonacci)

$$\mathrm{rep}_F(\mathbb{N}) = 1\{0, 01\}^* \cup \{\varepsilon\}$$

$\mathrm{rep}_F(n)$	n
ε	0
1	1
10	2
100	3
101	4
1000	5
1001	6
1010	7
10000	8

Decimation of languages

Let L be a language ordered w.r.t. the radix order.

If $w_0 < w_1 < \cdots$ are the elements of L and $X \subseteq \mathbb{N}$, then

$$L[X] = \{w_n : n \in X\}.$$

If $S = (L, \Sigma, <)$, then $L[X] = \mathrm{rep}_S(X)$.

If $L[X]$ is accepted by a finite automaton, what does it imply on X? What conditions on X insures that $L[X]$ is regular?

Motivation for ANS

ANS are a generalization of all usual PNS like integer base numeration systems or linear numeration systems, and even rational numeration systems.

Thanks to this general point of view on numeration systems, we try to distinguish results that deeply depend on the algorithm used to represent the integers from results that only depend on the set of representations.

Due to the general setting of ANS, some new questions concerning languages arise naturally from this numeration point of view.

Some questions around ANS

- Rec. sets in a given ANS?
- Rec. sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never recognizable?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is rec.?
- How do rec. depend on the choice of the numeration?
- For which ANS do arithmetic operations preserve rec.?
- Operations preserving rec. in a given ANS?
- How to represent real numbers?
- Can we define automatic sequences in that context?
- Logical characterization of rec. sets?
- Extensions to the multidimensional setting?
- ...

Some questions around ANS

- Rec. sets in a given ANS?
- Rec. sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never recognizable?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is rec.?
- How do rec. depend on the choice of the numeration?
- For which ANS do arithmetic operations preserve rec.?
- Operations preserving rec. in a given ANS?
- How to represent real numbers?
- Can we define automatic sequences in that context?
- Logical characterization of rec. sets?
- Extensions to the multidimensional setting?
- ...

Some questions around ANS

- Rec. sets in a given ANS?
- Rec. sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never recognizable?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is rec.?
- How do rec. depend on the choice of the numeration?
- For which ANS do arithmetic operations preserve rec.?
- Operations preserving rec. in a given ANS?
- How to represent real numbers?
- Can we define automatic sequences in that context?
- Logical characterization of rec. sets?
- Extensions to the multidimensional setting?
- ...

Some questions around ANS

- Rec. sets in a given ANS?
- Rec. sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never recognizable?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is rec.?
- How do rec. depend on the choice of the numeration?
- For which ANS do arithmetic operations preserve rec.?
- Operations preserving rec. in a given ANS?
- How to represent real numbers?
- Can we define automatic sequences in that context?
- Logical characterization of rec. sets?
- Extensions to the multidimensional setting?
- ...

Some questions around ANS

- Rec. sets in a given ANS?
- Rec. sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never recognizable?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is rec.?
- How do rec. depend on the choice of the numeration?
- For which ANS do arithmetic operations preserve rec.?
- Operations preserving rec. in a given ANS?
- How to represent real numbers?
- Can we define automatic sequences in that context?
- Logical characterization of rec. sets?
- Extensions to the multidimensional setting?
- ...

Some questions around ANS

- Rec. sets in a given ANS?
- Rec. sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never recognizable?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is rec.?
- How do rec. depend on the choice of the numeration?
- For which ANS do arithmetic operations preserve rec.?
- Operations preserving rec. in a given ANS?
- How to represent real numbers?
- Can we define automatic sequences in that context?
- Logical characterization of rec. sets?
- Extensions to the multidimensional setting?
- ...

Some questions around ANS

- Rec. sets in a given ANS?
- Rec. sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never recognizable?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is rec.?
- How do rec. depend on the choice of the numeration?
- For which ANS do arithmetic operations preserve rec.?
- Operations preserving rec. in a given ANS?
- How to represent real numbers?
- Can we define automatic sequences in that context?
- Logical characterization of rec. sets?
- Extensions to the multidimensional setting?
- ...

Some questions around ANS

- Rec. sets in a given ANS?
- Rec. sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never recognizable?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is rec.?
- How do rec. depend on the choice of the numeration?
- For which ANS do arithmetic operations preserve rec.?
- Operations preserving rec. in a given ANS?
- How to represent real numbers?
- Can we define automatic sequences in that context?
- Logical characterization of rec. sets?
- Extensions to the multidimensional setting?
- ...

Some questions around ANS

- Rec. sets in a given ANS?
- Rec. sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never recognizable?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is rec.?
- How do rec. depend on the choice of the numeration?
- For which ANS do arithmetic operations preserve rec.?
- Operations preserving rec. in a given ANS?
- How to represent real numbers?
- Can we define automatic sequences in that context?
- Logical characterization of rec. sets?
- Extensions to the multidimensional setting?
- ...

Some questions around ANS

- Rec. sets in a given ANS?
- Rec. sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never recognizable?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is rec.?
- How do rec. depend on the choice of the numeration?
- For which ANS do arithmetic operations preserve rec.?
- Operations preserving rec. in a given ANS?
- How to represent real numbers?
- Can we define automatic sequences in that context?
- Logical characterization of rec. sets?
- Extensions to the multidimensional setting?
- ...

Some questions around ANS

- Rec. sets in a given ANS?
- Rec. sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never recognizable?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is rec.?
- How do rec. depend on the choice of the numeration?
- For which ANS do arithmetic operations preserve rec.?
- Operations preserving rec. in a given ANS?
- How to represent real numbers?
- Can we define automatic sequences in that context?
- Logical characterization of rec. sets?
- Extensions to the multidimensional setting?
- ...

Some questions around ANS

- Rec. sets in a given ANS?
- Rec. sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never recognizable?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is rec.?
- How do rec. depend on the choice of the numeration?
- For which ANS do arithmetic operations preserve rec.?
- Operations preserving rec. in a given ANS?
- How to represent real numbers?
- Can we define automatic sequences in that context?
- Logical characterization of rec. sets?
- Extensions to the multidimensional setting?
- ...

S-automatic words

b-automatic words

An infinite word $x = (x_n)_{n\geq 0}$ is b-automatic if there exists a DFAO $\mathcal{A} = (Q, q_0, \Sigma_b, \delta, \Gamma, \tau)$ s.t. for all $n \geq 0$,

$$x_n = \tau(\delta(q_0, \mathrm{rep}_b(n))).$$

Theorem (Cobham 1972)

Let $b \geq 2$. An infinite word is b-automatic iff it is the image under a coding of an infinite fixed point of a b-uniform morphism.

S-automatic words

Let $S = (L, \Sigma, <)$ be an ANS.
An infinite word $x = (x_n)_{n\geq 0}$ is S-automatic if there exists a DFAO $\mathcal{A} = (Q, q_0, \Sigma, \delta, \Gamma, \tau)$ s.t. for all $n \geq 0$,

$$x_n = \tau(\delta(q_0, \mathrm{rep}_S(n))).$$

Theorem (Rigo-Maes 2002)

An infinite word is S-automatic for some ANS S iff it is the image under a coding of an infinite fixed point of a morphism, i.e. a morphic word.

Corollary

The set of primes is never S-recognizable.

Its characteristic sequence is not morphic (Mauduit 1988).

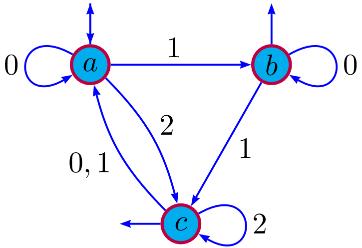
Corollary

The factor complexity of an S-automatic word is $O(n^2)$.

Idea of the proof

Example (Morphic → S-Automatic)

Consider the morphism μ defined by $a \mapsto abc$; $b \mapsto bc$; $c \mapsto aac$.
We have $\mu^\omega(a) = abcbcaacbcaacabcabcaacbcaacabcabc\cdots$.
One canonically associates the DFA $\mathcal{A}_{\mu,a}$



$$L_{\mu,a} = \{\varepsilon, 1, 2, 10, 11, 20, 21, 22, 100, 101, 110, 111, 112, 200, \ldots\}$$

If $S = (L_{\mu,a}, \{0,1,2\}, 0 < 1 < 2)$, then

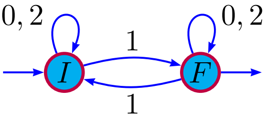
$$(\mu^\omega(a))_n = \delta_\mu(a, \mathrm{rep}_S(n)) \text{ for all } n \geq 0.$$

Idea of the proof

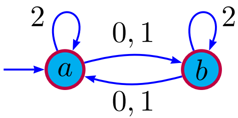
Example (S-Automatic $\to$ Morphic)

$S = (L, \{0,1,2\}, 0 < 1 < 2)$ where $L = \{w \in \Sigma^* : |w|_1 \text{ is odd}\}$

minimal automaton of L

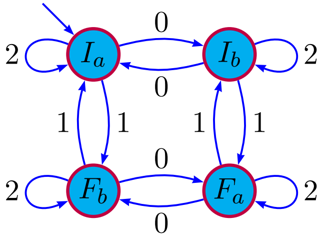


DFAO generating x



n	0	1	2	3	4	5	6	7	8	$\cdots$
$\text{rep}_S(n)$	1	01	10	12	21	001	010	012	021	$\cdots$
x	b	a	a	b	b	b	b	a	a	$\cdots$

Example (Continued)



$$
\begin{aligned}
f : \alpha &\mapsto \alpha I_b F_b I_a & F_a &\mapsto F_b I_b F_a & g : \alpha, I_a, I_b &\mapsto \varepsilon \\
I_a &\mapsto I_b F_b I_a & F_b &\mapsto F_a I_a F_b & F_a &\mapsto a \\
I_b &\mapsto I_a F_a I_b & & & F_b &\mapsto b
\end{aligned}
$$

$L \subseteq \Sigma^*$	ε	0	1	2	00	01	02	10	11	12	20
$f^\omega(\alpha)$	α	I_b	F_b	I_a	I_a	F_a	I_b	F_a	I_a	F_b	I_b
x			b			a		a		b	

$$g(f^\omega(\alpha)) = x$$

Multidimensional Case

A d-dimensional infinite word over an alphabet Σ is a map $x : \mathbb{N}^d \to \Sigma$. We use notation like $x_{n_1,\ldots,n_d}$ or $x(n_1, \ldots, n_d)$ to denote the value of x at $(n_1, \ldots, n_d)$.

If $w_1, \ldots, w_d$ are finite words over the alphabet Σ,

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

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

Example

$(ab, bbaa)^{\#} = (\#\#ab, bbaa) = (\#, b)(\#, b)(a, a)(b, a)$

A d-dimensional infinite word over an alphabet Γ is b-automatic if there exists a DFAO

$$\mathcal{A} = (Q, q_0, (\Sigma_b)^d, \delta, \Gamma, \tau)$$

s.t. for all $n_1, \ldots, n_d \geq 0$,

$$\tau\left(\delta\left(q_0, (\mathrm{rep}_b(n_1), \ldots, \mathrm{rep}_b(n_d))^0\right)\right) = x_{n_1,\ldots,n_d}.$$

Theorem (Salon 1987)

Let $b \geq 2$ and $d \geq 1$. A d-dimensional infinite word is b-automatic iff it is the image under a coding of a fixed point of a b-uniform d-dimensional morphism.

Theorem (C-Kärki-Rigo 2010)

Let $d \geq 1$. The d-dimensional infinite word is S-automatic for some ANS $S = (L, \Sigma, <)$ where $\varepsilon \in L$ iff it is the image under a coding of a shape-symmetric infinite d-dimensional word.

Shape-symmetric

$$\mu(a) = \mu(f) = \begin{array}{|c|c|}\hline a & b \\ \hline c & d \\ \hline\end{array}\,;\ \mu(b) = \begin{array}{|c|}\hline e \\ \hline c \\ \hline\end{array}\,;\ \mu(c) = \begin{array}{|c|c|}\hline e & b \\ \hline\end{array}\,;\mu(d) = \begin{array}{|c|}\hline f \\ \hline\end{array}$$

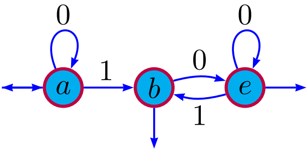
$$\mu(e) = \begin{array}{|c|c|}\hline e & b \\ \hline g & d \\ \hline\end{array}\,;\ \mu(g) = \begin{array}{|c|c|}\hline h & b \\ \hline\end{array}\,;\ \mu(h) = \begin{array}{|c|c|}\hline h & b \\ \hline c & d \\ \hline\end{array}\,.$$

$$\mu^{\omega}(a) = \begin{array}{|cc|c|cc|cc|c|c}\hline a & b & e & e & b & e & b & e & \cdots \\ c & d & c & g & d & g & d & c & \\ \hline e & b & f & e & b & h & b & f & \\ \hline e & b & e & a & b & e & b & e & \\ g & d & c & c & d & g & d & c & \\ \hline e & b & e & e & b & a & b & e & \\ g & d & c & g & d & c & d & c & \\ \hline h & b & f & e & b & e & b & f & \\ \hline \vdots & & & & & & & & \ddots \end{array}$$

Consider the morphism μ_1 defined by

$$a \mapsto ab\,;b \mapsto e\,;e \mapsto eb.$$

We have $\mu_1^\omega(a) = abeebebeebeebebeebebeebeeb\cdots$.
One canonically associates the DFA $\mathcal{A}_{\mu_1,a}$



$$L_{\mu_1,a} = \{\varepsilon, 1, 10, 100, 101, 1000, 1001, 1010, 10000, \ldots\}$$

Open question

- If S and T are two ANS, (S,T)-automatic words are bidimensional infinite words $(x_{m,n})_{m,n\geq 0}$ for which there exists a DFAO $\mathcal{A} = (Q, (\Sigma \cup \{\#\})^d, \delta, q_0, \Gamma, \tau)$ s.t. $\forall m, n \in \mathbb{N}$,

$$x_{m,n} = \tau(\delta(q_0, (\mathrm{rep}_S(m), \mathrm{rep}_T(n))^{\#})).$$

  Can these (S,T)-automatic words be characterized by iterating morphisms?

b-kernel

An infinite word $(x_n)_{n\geq 0}$ is b-automatic iff its b-kernel

$$\{(x_{b^e n+r})_{n\geq 0} \colon e, r \in \mathbb{N},\ r < b^e\}$$

is finite. The b-kernel can be rewritten

$$\{(x_{b^{|w|}n+\mathrm{val}_b(w)})_{n\geq 0} \colon w \in \Sigma_b^*\}.$$

8	4	2	1		8	4	2	1			8	4	2	1	
			ε	0		1	1	0	6		1	1	0	0	12
			1	1		1	1	1	7		1	1	0	1	13
		1	0	2	1	0	0	0	8		1	1	1	0	14
		1	1	3	1	0	0	1	9		1	1	1	1	15
	1	0	0	4	1	0	1	0	10	1	0	0	0	0	16
	1	0	1	5	1	0	1	1	11	1	0	0	0	1	17

NB: $b^{|w|}n + \mathrm{val}_b(w)$ is the base-b value of the $(n+1)$-th word in $\mathcal{L}_b$ having w as a suffix.

Open question

The S-kernel of $(x_n)_{n\geq 0}$ is

$$\{(x_{f_w(n)})_{n\geq 0} \colon w \in \Sigma^*\}$$

where $f_w(n)$ is the S-value of the $(n+1)$-th word in L having w as a suffix.

Theorem (Rigo-Maes 2002)

An infinite word is S-automatic iff its S-kernel is finite.

- Does a similar characterization hold in the multidimensional setting?

Sets S-recognizable for all S

Ultimately periodic sets

It is an exercise to show that all ultimately periodic set are b-recognizable for all $b \geq 2$.

Theorem (Cobham 1969)

Let $k, \ell \geq 2$ be two multiplicatively independent integers. A subset of $\mathbb{N}$ is both k-recognizable and ℓ-recognizable iff it is ultimately periodic.

Two numbers k and ℓ are multiplicatively independent if $k^m = \ell^n$ and $m, n \in \mathbb{N}$ implies $m = n = 0$.

Corollary

A subset of $\mathbb{N}$ is b-recognizable for all $b \geq 2$ iff it is ultimately periodic.

Generalization to ANS

Theorem (Lecomte-Rigo 2001, Krieger et al. 2009)

Ultimately periodic sets are S-recognizable for all ANS S.

Corollary

A subset of $\mathbb{N}$ is S-recognizable for all ANS S iff it is ultimately periodic.

Theorem (Krieger et al. 2009, Angrand-Sakarovitch 2010)

Let $m, r \in \mathbb{N}$ with $m \geq 2$ and $0 \leq r \leq m-1$ and let $S = (L, \Sigma, <)$ be an ANS. If L is accepted by a n-state DFA, then the minimal DFA of $\mathrm{rep}_S(m\mathbb{N} + r)$ has at most nm^n states.

Semi-linear sets

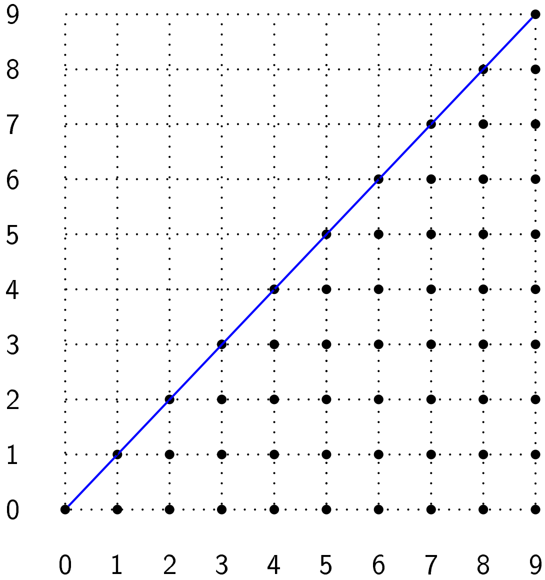
A subset X of $\mathbb{N}^d$ is b-recognizable if the language $(\mathrm{rep}_b(X))^{\#}$ over $(\{0,1,\ldots,b-1\}\cup\{\#\})^d$ is regular, where

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

Theorem (Cobham–Semenov, Semenov 1977)

Let $k,\ell\geq 2$ be two multiplicatively independent integers. A subset of $\mathbb{N}^d$ is both k-recognizable and ℓ-recognizable iff it is semi-linear.

A set $X\subseteq\mathbb{N}^d$ is linear if there exist $\mathbf{v}_0,\mathbf{v}_1,\cdots,\mathbf{v}_t\in\mathbb{N}^d$ such that $X=\mathbf{v}_0+\mathbb{N}\,\mathbf{v}_1+\mathbb{N}\,\mathbf{v}_2+\cdots+\mathbb{N}\,\mathbf{v}_t$. A set $X\subseteq\mathbb{N}^d$ is semi-linear if it is a finite union of linear sets.



$$\{(n,m) : n,m \in \mathbb{N} \text{ and } n \geq m\} = \mathbb{N}(1,0) + \mathbb{N}(1,1)$$

Semi-linear sets: a good generalization?

Corollary

A subset of $\mathbb{N}^d$ is b-recognizable for all $b \geq 2$ iff it is semi-linear.

In the one-dimensional case, we have the following equivalences:

semi-linear $\Leftrightarrow$ ultimately periodic $\Leftrightarrow$ 1-recognizable.

Multidimensional case for ANS

One might therefore expect that the semi-linear sets are recognizable in all ANS. However, this fails to be the case, as the following example shows.

Example

The semi-linear set $X = \{n(1,2) : n \in \mathbb{N}\} = \{(n, 2n) \mid n \in \mathbb{N}\}$ is not 1-recognizable. Consider the language $\{(a^n\#^n, a^{2n}) \mid n \in \mathbb{N}\}$, consisting of the unary representations of the elements of X.
Use the pumping lemma to show that this is not accepted by a finite automaton.

Let $S = (L, \Sigma, <)$ be an ANS.

A subset X of $\mathbb{N}^d$ is S-recognizable if the language $(\mathrm{rep}_S(X))^{\#}$ over $(\Sigma \cup \{\#\})^d$ is regular, where

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

It is 1-recognizable if it is S-automatic for the ANS S built on a^*.

Multidimensional 1-recognizable sets

Theorem (C-Lacroix-Rampersad 2012)

A subset of $\mathbb{N}^d$ is S-recognizable for all ANS S iff it is 1-recognizable.

Theorem (C-Lacroix-Rampersad 2012)

The multidimensional 1-recognizable sets are the finite unions of sets of the form

$$(a_1 + b_1\,\mathbb{N})\mathbf{v}_1 + \cdots + (a_t + b_t\,\mathbb{N})\mathbf{v}_t,$$

where

- $\mathrm{Supp}(\mathbf{v}_1) \supseteq \mathrm{Supp}(\mathbf{v}_2) \supseteq \cdots \supseteq \mathrm{Supp}(\mathbf{v}_t)$
- *All $\mathbf{v}_i$ are vectors all of whose components are 0 or 1.*

Recognizable 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 the right congruence $\sim_X$ has finite index ($x \sim_X y$ if $\forall z \in \mathbb{N}^d \ (x + z \in X \Leftrightarrow y + z \in X)$).

When $d = 1$, we have again the following equivalences:

recognizable $\Leftrightarrow$ ultimately periodic $\Leftrightarrow$ 1-recognizable.

However, for $d > 1$ these equivalences no longer hold.

Multidimensional recognizable sets: a characterization

Theorem (Mezei)

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 $D = \{(n, n) \mid n \in \mathbb{N}\}$ is not recognizable.

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.