S-automatic sets

Émilie Charlier

University of Waterloo, School of Computer Science

Séminaire du Lacim, Montréal, December 3, 2010

Numeration systems

A numeration system is a bijection between $\mathbb{N}$ and a language:

$$\mathrm{rep}\colon \mathbb{N} \to L.$$

Common desired properties:

- The numeration language L is accepted by a finite automaton;
- The growth functions of L are exponential;
- The growth functions of L satisfy some linear recurrences;
- Targeted subsets X of $\mathbb{N}$ have simple sets of representations $\mathrm{rep}(X)$;
- The representations of any arithmetic progression form a language accepted by a finite automaton;
- ...

Abstract numeration systems

An *abstract numeration system* (ANS) is a triple $S = (L, \Sigma, <)$ where L is an infinite language over a totally ordered 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-automatic* if $\mathrm{rep}_S(X)$ is accepted by a finite automaton.

1-automatic sets

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

n	0	1	2	3	4	5	$\cdots$
$\mathrm{rep}_1(n)$	ε	a	aa	aaa	$aaaa$	$aaaaa$	$\cdots$

A subset X of $\mathbb{N}$ is *1-automatic* if it is S-automatic for the ANS S built on a^*.

Theorem (Characterization of 1-automatic sets)

A subset of $\mathbb{N}$ is 1-automatic iff it is ultimately periodic.

Second example

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

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

For instance, the sets

$$\mathrm{val}_S(a^*) = \{2^n - 1 \mid n \in \mathbb{N}\}$$

and

$$\mathrm{val}_S(a^*b^*) = \{1, 2, 3, 4, 6, 7, 8, 10, 14, 15, \ldots\}$$

are S-automatic.

Third example: a polynomial numeration language

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

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

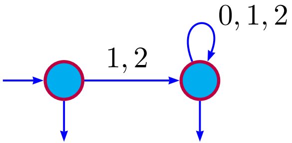
The set

$$\mathrm{val}_S(a^*) = \{\frac{(n+2)(n+1)}{2} \mid n \in \mathbb{N}\}$$

is S-automatic.

No set of integers whose n-th term grows faster than a polynomial can be S-automatic.

Integer base $b \geq 2$ numeration system

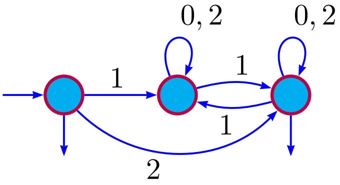


$\mathrm{rep}_3(\mathbb{N})$

$\mathbb{N}$ is 3-automatic

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 $b \geq 2$ numeration system



$\mathrm{rep}_3(2\,\mathbb{N})$

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

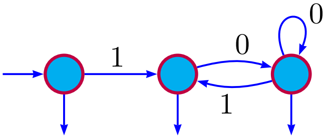
27	9	3	1	
			ε	0
			1	1
			2	2
		1	0	3
		1	1	4
		1	2	5
		2	0	6
		2	1	7
		2	2	8
	1	0	0	9

Fibonacci numeration system

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

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

The factor 11 is forbidden:



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

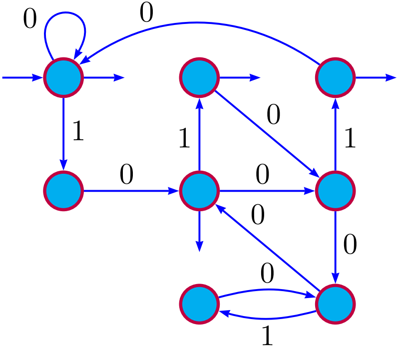
$\mathbb{N}$ is F-automatic

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 numeration system

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

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



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

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

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

Automaticity of $\mathbb{N}$

Is the set $\mathbb{N}$ automatic? Otherwise stated, is the *numeration language* accepted by a finite automaton? Not necessarily:

Theorem (Shallit 1994)

Let U be a sequence generating a (positional) numeration system. If $\mathbb{N}$ is automatic, then U is linear, that is, 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".

Motivation for ANS

ANS are a generalization of all usual (positional) numeration systems 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

- Automatic sets in a given ANS?
- Automatic sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never automatic?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is automatic?
- For which ANS do arithmetic operations preserve automaticity?
- What operations do preserve automaticity in a given ANS?
- How to represent real numbers ?
- Can we define automatic sequences in that context?
- Logical characterization of automatic sets?
- ...

Some questions around ANS

- Automatic sets in a given ANS?
- Automatic sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never automatic?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is automatic?
- For which ANS do arithmetic operations preserve automaticity?
- What operations do preserve automaticity in a given ANS?
- How to represent real numbers ?
- Can we define automatic sequences in that context?
- Logical characterization of automatic sets?
- ...

Some questions around ANS

- Automatic sets in a given ANS?
- Automatic sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never automatic?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is automatic?
- For which ANS do arithmetic operations preserve automaticity?
- What operations do preserve automaticity in a given ANS?
- How to represent real numbers ?
- Can we define automatic sequences in that context?
- Logical characterization of automatic sets?
- ...

Some questions around ANS

- Automatic sets in a given ANS?
- Automatic sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never automatic?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is automatic?
- For which ANS do arithmetic operations preserve automaticity?
- What operations do preserve automaticity in a given ANS?
- How to represent real numbers ?
- Can we define automatic sequences in that context?
- Logical characterization of automatic sets?
- ...

Some questions around ANS

- Automatic sets in a given ANS?
- Automatic sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never automatic?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is automatic?
- For which ANS do arithmetic operations preserve automaticity?
- What operations do preserve automaticity in a given ANS?
- How to represent real numbers ?
- Can we define automatic sequences in that context?
- Logical characterization of automatic sets?
- ...

Some questions around ANS

- Automatic sets in a given ANS?
- Automatic sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never automatic?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is automatic?
- For which ANS do arithmetic operations preserve automaticity?
- What operations do preserve automaticity in a given ANS?
- How to represent real numbers ?
- Can we define automatic sequences in that context?
- Logical characterization of automatic sets?
- ...

Some questions around ANS

- Automatic sets in a given ANS?
- Automatic sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never automatic?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is automatic?
- For which ANS do arithmetic operations preserve automaticity?
- What operations do preserve automaticity in a given ANS?
- How to represent real numbers ?
- Can we define automatic sequences in that context?
- Logical characterization of automatic sets?
- ...

Some questions around ANS

- Automatic sets in a given ANS?
- Automatic sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never automatic?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is automatic?
- For which ANS do arithmetic operations preserve automaticity?
- What operations do preserve automaticity in a given ANS?
- How to represent real numbers ?
- Can we define automatic sequences in that context?
- Logical characterization of automatic sets?
- ...

Some questions around ANS

- Automatic sets in a given ANS?
- Automatic sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never automatic?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is automatic?
- For which ANS do arithmetic operations preserve automaticity?
- What operations do preserve automaticity in a given ANS?
- How to represent real numbers ?
- Can we define automatic sequences in that context?
- Logical characterization of automatic sets?
- ...

Some questions around ANS

- Automatic sets in a given ANS?
- Automatic sets in all ANS?
- Are there subsets of $\mathbb{N}$ that are never automatic?
- Given a subset of $\mathbb{N}$ can we build an ANS for which it is automatic?
- For which ANS do arithmetic operations preserve automaticity?
- What operations do preserve automaticity in a given ANS?
- How to represent real numbers ?
- Can we define automatic sequences in that context?
- Logical characterization of automatic sets?
- ...

First part

Ultimately periodic sets

It is an exercise to show that any ultimately periodic set is b-automatic for all $b \geq 1$.

Theorem (Cobham 1969)

Let $k, \ell \geq 2$ be two multiplicatively independent integers. A subset of $\mathbb{N}$ is both k-automatic and ℓ-automatic 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 $m = n = 0$.

b-automatic sets for all $b \geq 2$

Corollary

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

Another formulation of this result is:

Corollary

A subset of $\mathbb{N}$ *is* b*-automatic for all* $b \geq 2$ *iff it is* 1*-automatic.*

Generalization to ANS

Theorem (Lecomte, Rigo 2001)

Ultimately periodic sets are S-automatic for all ANS S built on a regular language.

Corollary

A subset of $\mathbb{N}$ is S-automatic for all ANS S built on a regular language iff it is ultimately periodic.

Corollary

A subset of $\mathbb{N}$ is S-automatic for all ANS S built on a regular language iff it is 1-automatic.

Multi-dimensional case: $\mathbb{N}^d,\ d \geq 1$

A subset X of $\mathbb{N}^d$ is *b-automatic* if the language $(\mathrm{rep}_b(X))^\#$ over $(\{0,1,\ldots,b-1\} \cup \{\#\})^d$ is accepted by a finite automaton, where $\mathrm{rep}_b(X) = \{(\mathrm{rep}_b(n_1),\ldots,\mathrm{rep}_b(n_d))\colon (n_1,\ldots,n_d) \in X\}$ and where $(\cdot)^\#$ is the padding map.

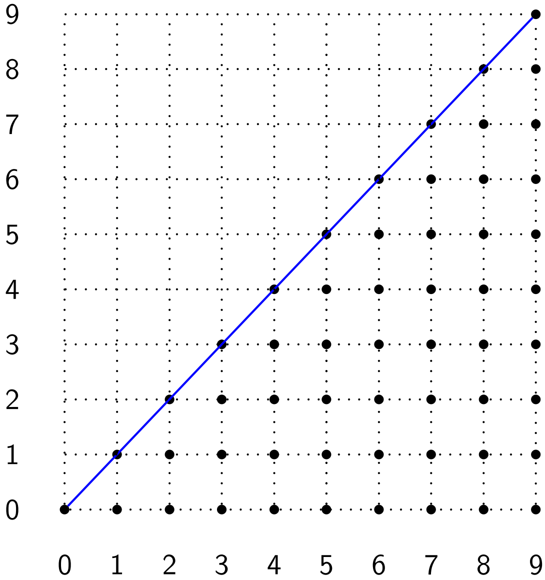
$$(aba,a)^\# = (aba,a\#\#) = (a,a)(b,\#)(a,\#) \in (\{a,b,\#\}^2)^*$$

Semi-linear sets

Theorem (Cobham–Semenov, Semenov 1977)

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

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

$$\text{semi-linear} \Leftrightarrow \text{ultimately periodic} \Leftrightarrow 1\text{-automatic}.$$

So, semi-linear sets seem to be a good generalization of ultimately periodic sets.

Multi-dimensional case for ANS

One might therefore expect that the semi-linear sets are automatic 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-automatic. 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.

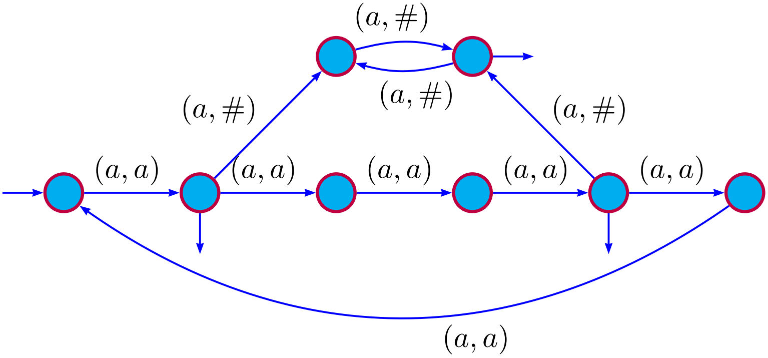
A subset X of $\mathbb{N}^d$ is *S-automatic* if the language $(\mathrm{rep}_S(X))^{\#}$ over $(\Sigma \cup \{\#\}))^d$ is accepted by a finite automaton,
where $\mathrm{rep}_S(X) = \{(\mathrm{rep}_S(n_1), \ldots, \mathrm{rep}_S(n_d)) \colon (n_1, \ldots, n_d) \in X\}$.

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

Example

Let $X = \{(2n, 3m+1) \mid n, m \in \mathbb{N}$ and $2n \geq 3m+1\}$.

It is clear that X is 1-automatic:



Example (cont'd)

Let $X = \{(2n, 3m+1) \mid n, m \in \mathbb{N} \text{ and } 2n \geq 3m+1\}$.

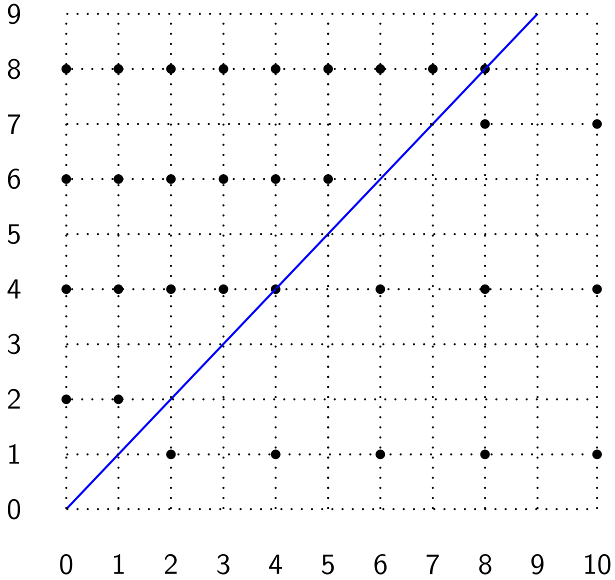
Let S be an ANS.

The sets $2\,\mathbb{N}$ and $3\,\mathbb{N}+1$ are both S-automatic, and the set $(2\,\mathbb{N} \times (3\,\mathbb{N}+1))^{\#}$ is accepted by a finite automaton.

Furthermore, the set $\{(x, y)^{\#} \mid x, y \in L \text{ and } x \geq y\}$ is also accepted by a finite automaton.

By taking the product of these two automata we obtain an automaton accepting
$\{(\mathrm{rep}_S(2n), \mathrm{rep}_S(3m+1))^{\#} \mid n, m \in \mathbb{N} \text{ and } 2n \geq 3m+1\}$.

Then we see that X is S-automatic.



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

$$\cup \{(n, 2m) : n, m \in \mathbb{N} \text{ and } n < 2m\}$$

Theorem (C, Lacroix, Rampersad 2010)

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

In view of this theorem, the correct generalization of ultimately periodic sets for ANS turns out to be 1-automatic sets.

Multi-dimensional 1-automatic sets

Let $\emptyset \neq A \subseteq \{1, \ldots, d\}$. Define Σ_A to be the subalphabet $\{x \in (\Sigma_\#)^d \mid \text{the } i\text{-th component of } x \text{ is } \# \text{ exactly when } i \notin A\}$ where $\Sigma_\# = \Sigma \cup \{\#\}$.

Example

Let $\Sigma = \{a\}$ and $d = 4$.
If $A = \{1, 2, 3, 4\}$, then $\Sigma_A = \{(a, a, a, a)\}$.
If $A = \{2, 3\}$, then $\Sigma_A = \{(\#, a, a, \#)\}$.
If $A = \{3\}$, then $\Sigma_A = \{(\#, \#, a, \#)\}$.

Theorem (Decomposition, Eilenberg, Elgot, Shepherdson 1969)

Let $R \subseteq (\Sigma^*)^d$. The language $R^\# \subseteq ((\Sigma_\#)^d)^*$ is regular iff it is a finite union of languages of the form $R_0 \cdots R_t$, where $t \in \mathbb{N}$, each factor $R_i \subseteq (\Sigma_{A_i})^*$ is regular, and $A_t \subseteq \cdots \subseteq A_0 \subseteq \{1, \ldots, d\}$.

Example

Let X be the set

$$\{(5n, 5n+4m+6\ell+1, 5n+4m+6\ell+3, 5n) : n, m, \ell \in \mathbb{N}\}.$$

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-automatic.

The set X can be written as

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

Example

Let $R = \{(a^{5n}, a^{6m}) \mid n, m \in \mathbb{N}\}$. Then $R^{\#}$ is regular and

$$R^{\#} = \bigcup_{\ell=0}^{5} (a^{30}, a^{30})^* (a^{5\ell} \#^{\ell}, a^{6\ell})(\#^6, a^6)^*$$

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

The corresponding set of integers $X = \{5n, 6m) \mid n, m \in \mathbb{N}\}$ can be written as

$$X = \bigcup_{\ell=0}^{5} \{30(n,n) + (5\ell, 5\ell) + (0, \ell) + 6(0, m) \mid n, m \in \mathbb{N}\}$$

$$\cup \bigcup_{\ell=0}^{4} \{30(n,n) + (6\ell, 6\ell) + (5-\ell, 0) + 5(m, 0) \mid n, m \in \mathbb{N}\}.$$

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-automatic.

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

Multi-dimensional 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-automatic subset of $\mathbb{N}^2$.

So we see that for $d > 1$, the class of 1-automatic sets corresponds neither to the class of semi-linear sets, nor to the class of recognizable sets.

Second part

Which growth functions for automatic sets?

Let L be a language over an alphabet Σ. Define

$$\mathbf{u}_L(n) = |L \cap \Sigma^n| \quad \text{and} \quad \mathbf{v}_L(n) = \sum_{i=0}^{n} \mathbf{u}_L(i) = |L \cap \Sigma^{\leq n}|$$

The maps $\mathbf{u}_L \colon \mathbb{N} \to \mathbb{N}$ and $\mathbf{v}_L \colon \mathbb{N} \to \mathbb{N}$ are the *growth functions* of L.

If $X \subseteq \mathbb{N}$, we let $t_X(n)$ denote the $(n+1)$-th term of X.
The map $t_X \colon \mathbb{N} \to \mathbb{N}$ is the *growth function* of X.

What do the growth functions of S-automatic sets look like?

Another formulation

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

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

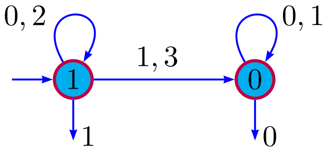
$$L_X = \{w_n \mid n \in X\}.$$

If L_X is accepted by a finite automaton, what does it imply on X?

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

Example

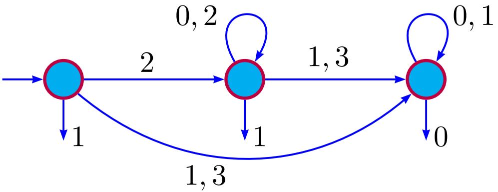
Consider the morphism h defined by $h(1) = 1010$ and $h(0) = 00$.



$h^\omega(1)$	1	0	1	0	0	0	1	0	1	0	0	$\cdots$
$X \subseteq \mathbb{N}$	0	1	2	3	4	5	6	7	8	9	10	$\cdots$

Example (cont'd)

The set X is S-automatic for the ANS S built on the language accepted by the DFA:



$h^\omega(1)$	1	0	1	0	0	0	1	0	1	0	0	0	$\cdots$
$X \subseteq \mathbb{N}$	0	1	2	3	4	5	6	7	8	9	10	11	$\cdots$
L	ε	1	2	3	10	11	20	21	22	23	30	31	$\cdots$

$$\mathrm{rep}_S(X) = 2\{0,2\}^* \cup \{\varepsilon\}$$

Example (cont'd)

For all $k \in \mathbb{N}$, we have $|h^k(1)| = (k+1)2^k$.

For all $n \in \mathbb{N}$, there is a unique $k := k(n) \in \mathbb{N}$ such that $(k+1)2^k \leq t_X(n) < (k+2)2^{k+1}$.

Since the number of occurrences of 1 in the prefix $h^k(1)$ is 2^k, we also have $(k+1)2^k \leq t_X(n) < (k+2)2^{k+1} \Leftrightarrow 2^k \leq n < 2^{k+1}$.

This means that $k(n) = \log_2(n)$. Consequently, $t_X(n)$ is $\Theta(n \log(n))$.

$\mathbf{v}_L(k) = (k+1)2^k$ and $\mathbf{v}_{\mathrm{rep}_S(X)}(k) = 2^k$

Theorem (C, Rampersad 2010)

Let $S = (L, \Sigma, <)$ be an ANS built on a regular language and let $X \subseteq \mathbb{N}$ be an infinite S-automatic set. Suppose

$$\forall i \in \{0, \ldots, p-1\},\ \mathbf{v}_L(np+i) \sim a_i n^c \alpha^n\ (n \to +\infty),$$

for some $p, c \in \mathbb{N}$ with $p \geq 1$, some $\alpha \geq 1$ and some positive constants $a_0, \ldots, a_{p-1}$, and

$$\forall j \in \{0, \ldots, q-1\},\ \mathbf{v}_{\mathrm{rep}_S(X)}(nq+j) \sim b_j n^d \beta^n\ (n \to +\infty),$$

for some $q, d \in \mathbb{N}$ with $q \geq 1$, some $\beta \geq 1$ and some positive constants $b_0, \ldots, b_{q-1}$. Then we have

- $t_X(n) = \Theta\left((\log(n))^{c - d\frac{\log(\sqrt[p]{\alpha})}{\log(\sqrt[q]{\beta})}} n^{\frac{\log(\sqrt[p]{\alpha})}{\log(\sqrt[q]{\beta})}} \right)$ *if $\beta > 1$;*
- $t_X(n) = \Theta\left(n^{\frac{c}{d}} (\sqrt[p]{\alpha})^{\Theta(n^{\frac{1}{d}})} \right)$ *if $\beta = 1$.*

Background on S-automatic sequences

Let S be an ANS. An infinite word x over an alphabet Γ is *S-automatic* if, for all $n \in \mathbb{N}$, its $(n+1)$st letter $x[n]$ is obtained by "feeding" a DFAO $\mathcal{A} = (Q, \Sigma, \delta, q_0, \Gamma, \tau)$ with the S-representation of n: $\forall n \in \mathbb{N},\ \tau(\delta(q_0, \mathrm{rep}_S(n))) = x[n]$.

Let $X \in \mathbb{N}$. Its *characteristic word* is the infinite word $(\chi_X[n])_{n \geq 0}$ defined by

$$\chi_X[n] = \begin{cases} 1, & \text{if } n \in X; \\ 0, & \text{otherwise.} \end{cases}$$

Proposition

Let S be an ANS. A set is S-automatic iff its characteristic word is S-automatic.

Background on S-automatic sequences

If μ is a morphism over an alphabet Σ and a is a letter in Σ such that the image $\mu(a)$ begins with a, then μ is *prolongable on a*.

If a morphism μ is prolongable on a letter a, then the limit $\mu^\omega(a) := \lim_{n\to+\infty} \mu^n(a)$ is well defined.

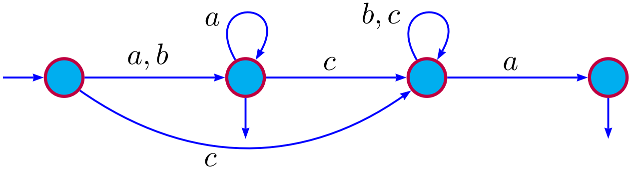
An infinite word is *pure morphic* if it can be written as $\mu^\omega(a)$ for some morphism μ prolongable on a. It is *morphic* if it the image under a morphism of some pure morphic word.

Theorem (Rigo 2000, Rigo, Maes 2002)

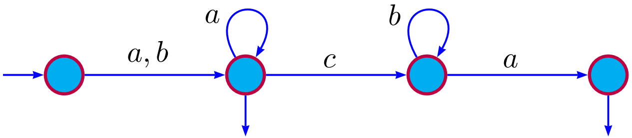
An infinite word is S-automatic for some ANS S if and only if it is morphic.

Proof on an example

Let $S = (L, \{a, b, c\}, a < b < c)$ be an ANS where L is the language accepted by the (trim) DFA $\mathcal{A}_L$:



Let $X = \mathrm{val}_S(M)$ where $M \subseteq L$ is the language accepted by the following (trim) DFA $\mathcal{A}_X$:



Clearly, X is S-automatic.

Proof on an example (cont'd)

We have $X = \{0, 1, 2, 3, 5, 6, 7, 8, 11, 12, 13, 15, 16, 17, 23, \ldots\}$.

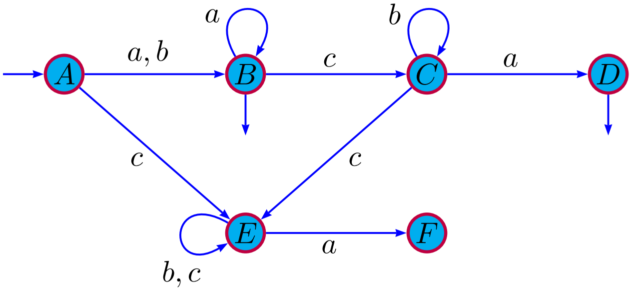
For all $k \geq 1$, we have

$$\mathbf{v}_L(k) = 3 \cdot 2^{k-1} - 1 \quad \text{and} \quad \mathbf{v}_{\mathrm{rep}_S(X)}(k) = \mathbf{v}_M(k) = k^2 - k + 2.$$

.

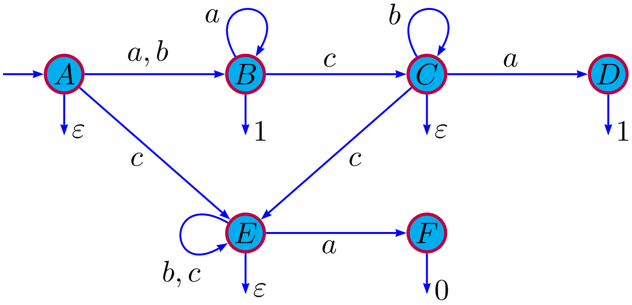
Proof on an example (cont'd)

Consider the accessible part of the product automaton of $\mathcal{A}_L$ and $\mathcal{A}_X$ (the latter has been completed first):

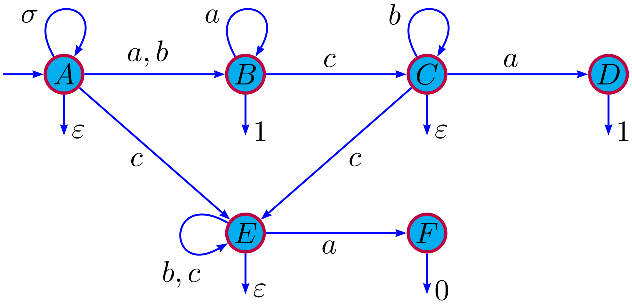


Proof on an example (cont'd)

Consider the accessible part of the product automaton of $\mathcal{A}_L$ and $\mathcal{A}_X$ (the latter has been completed first):



Proof on an example (cont'd)



Suppose $\sigma < a < b < c$. We obtain the morphisms

$$
\begin{array}{rlrlr}
h\colon A \mapsto ABBE & & D \mapsto \varepsilon & & g\colon A, C, E \mapsto \varepsilon \\
B \mapsto BC & & E \mapsto FEE & & B, D \mapsto 1 \\
C \mapsto DCE & & F \mapsto \varepsilon & & F \mapsto 0
\end{array}
$$

Let K be the language accepted by the underlying DFA where all states are final.

$L \subseteq K$	ε	a	b	c	aa	ac	ba	bc	ca	cb	cc	aaa
$h^\omega(A)$	A	B	B	E	B	C	B	C	F	E	E	B
$g(h^\omega(A))$		1	1		1		1		0		1	

aac	aca	acb	acc	baa	bac	bca	bcb	bcc	cba	...
C	D	C	E	B	C	D	C	E	F	...
	1			1		1			0	...

$X = \{0, 1, 2, 3, 5, 6, 7, 8, 11, 12, 13, 15, 16, 17, 23, \dots\}$

$\mathbf{v}_K(k) = |h^k(A)|$, $\mathbf{v}_L(k) = |g(h^k(A))|$, and $\mathbf{v}_{\mathrm{rep}_S(X)}(k)$ is equal to the number $F(k)$ of 1's in $g(h^k(A))$.

$|g(h^k(A))| \leq t_X(n) < |g(h^{k+1}(A))| \Leftrightarrow F(k) \leq n < F(k+1)$.

Since $\mathbf{v}_L(k) = 3 \cdot 2^{k-1} - 1$ and $\mathbf{v}_{\mathrm{rep}_S(X)}(k) = k^2 - k + 2$, we obtain

$$t_X(n) = 2^{\Theta(\sqrt{n})}.$$

Exponential numeration language

Proposition (C, Rampersad, 2010)

For every integers $k \geq 0$ and $\ell \geq 1$, there exists an ANS S built on an exponential regular language and an infinite S-automatic set $X \subseteq \mathbb{N}$ such that $t_X(n) = \Theta((\log(n))^k n^\ell)$.

Choose $\alpha = \beta^\ell$.

Polynomial numeration language

Corollary

Let $S = (L, \Sigma, <)$ *be an ANS built on a polynomial regular language and let* $X \subseteq \mathbb{N}$ *be an infinite* S*-automatic set. Then we have* $t_X(n) = \Theta(n^r)$ *for some rational number* $r \geq 1$*.*

Proposition (C, Rampersad, 2010)

For every rational number $r \geq 1$*, there exists an ANS* S *built on a polynomial regular language and an infinite* S*-automatic set* $X \subseteq \mathbb{N}$ *such that* $t_X(n) = \Theta(n^r)$*.*

An example

Consider the ANS $S = (a^*b^*c^*, \{a, b, c\}, a < b < c)$.

Let $X = \mathrm{val}_S(a^*c^*) = \{0, 1, 3, 4, 6, 9, 10, 12, 15, 19, 20, \ldots\}$. It is clearly S-automatic.

We have $\mathbf{v}_{a^*b^*c^*}(n) = \binom{n+3}{3}$ and $\mathbf{v}_{a^*b^*}(n) = \binom{n+2}{2}$.

We obtain

$$t_X(n) = \Theta(n^{\frac{3}{2}}).$$

Proposition

The set of squares $\{n^2 \mid n \in \mathbb{N}\}$ *is not* b*-automatic for all integer bases* $b \geq 2$*.*

But it is S-automatic for the ANS built on $a^*b^* \cup a^*c^*$ with $a < b < c$ since we have $\mathrm{rep}_S(\{n^2 \mid n \in \mathbb{N}\}) = a^*$.

Proposition (Rigo, 2002)

For all $k \in \mathbb{N}$*, the set* $\{n^k \mid n \in \mathbb{N}\}$ *is* S*-automatic for some* S*.*

In those constructions, the exhibited ANS are built on polynomial languages.

An example

Consider the base 4 numeration system, that is, the ANS built on $\mathcal{L}_4 = \{\varepsilon\} \cup \{1,2,3\}\{0,1,2,3\}^*$ with the natural order on the digits.

Let $X = \mathrm{val}_4(\{1,3\}^*) = \{1,3,5,7,13,15,21,23,29,31,\ldots\}$. It is clearly 4-automatic.

We have $\mathbf{v}_{\mathcal{L}_4}(n) = 4^n$ and $\mathbf{v}_{\{1,3\}^*}(n) = 2^{n+1} - 1$.

We obtain

$$t_X(n) = \Theta(n^2).$$

Theorem (Durand, Rigo 2009)

Let S be an ANS built on a polynomial regular language and let T be an ANS built on an exponential regular language. If a subset of $\mathbb{N}$ is both S-automatic and T-automatic, then it is ultimately periodic.

Corollary

Let S be an ANS built on an exponential regular language. If $f \in \mathbb{Q}[x]$ is a polynomial of degree greater than 1 such that $f(\mathbb{N}) \subseteq \mathbb{N}$, then the set $f(\mathbb{N})$ is not S-automatic.

Additional examples

Consider the 4-automatic set
$Y = \mathrm{val}_4(\{1,2,3\}^*) = \{0,1,3,5,6,7,9,10,11,13,14,15,21,\ldots\}$.

We have $\mathbf{v}_{\{1,2,3\}^*}(n) = \frac{1}{2}(3^{n+1} - 1)$.

We obtain

$$t_Y(n) = \Theta\left(n^{\frac{\log(4)}{log(3)}}\right).$$

Additional examples

Define $L_F = \{\varepsilon\} \cup 1(0+01)^*$ be the Fibonacci language and consider the 4-automatic $Z = \mathrm{val}_4(L_F)$. We have $Z = \{0, 1, 4, 16, 17, 64, 70, 256, 257, 260, 272, 273, 1024, \ldots\}$.

We have

$$\mathbf{v}_{L_F}(n) \sim \frac{5+3\sqrt{5}}{10}\left(\frac{1+\sqrt{5}}{2}\right)^n \quad (n \to +\infty).$$

We obtain

$$t_Z(n) = \Theta\left(n^{\frac{\log(4)}{\log(\varphi)}}\right)$$

where $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden ratio.