Syntactic complexity of ultimately periodic sets of integers

Michel Rigo, Elise Vandomme

Numeration 2011 – Liège

Université de Liège

GRENOBLE 1 UNIVERSITE JOSEPH FOURIER SCIENCES.TECHNOLOGIE.SANTÉ

Baum-Sweet sequence

Let $A = \{a, b, c, d\}$, $B = \{0, 1\}$,

$$f: \begin{array}{l} a \mapsto ab \\ b \mapsto cb \\ c \mapsto bd \\ d \mapsto dd \end{array} \quad \text{and} \quad g: \begin{array}{l} a \mapsto 1 \\ b \mapsto 1 \\ c \mapsto 0 \\ d \mapsto 0 \end{array}$$

Baum-Sweet sequence

Let $A = \{a, b, c, d\}$, $B = \{0, 1\}$,

$$f: \begin{array}{l} a \mapsto ab \\ b \mapsto cb \\ c \mapsto bd \\ d \mapsto dd \end{array} \quad \text{and} \quad g: \begin{array}{l} a \mapsto 1 \\ b \mapsto 1 \\ c \mapsto 0 \\ d \mapsto 0 \end{array}$$

$$\begin{array}{l} a \\ \downarrow f \\ ab \\ \downarrow f \\ abcb \\ \downarrow f \\ abcbbdcb \\ \downarrow f \\ \vdots \end{array}$$

We have

$$f^\omega(a) = abcbbdcbcbddbd \cdots$$

$$\begin{aligned} (x_n)_{n \geq 0} &:= g(f^\omega(a)) \\ &= 11011001010010 \ldots \end{aligned}$$

Periodicity problem

Let A, B be finite alphabets.

A morphism $f : A \to B$ is prolongable on $a \in A$ if

$$f(a) = aw \text{ with } w \in A^* \setminus \{\varepsilon\}.$$

A coding $g : A \to B$ is a letter-to-letter morphism.

General problem (HD0L periodicity problem)

Let

- $g : A \to B$ be a coding ,
- $f : A \to A^*$ be a morphism prolongable on $a \in A^*$.

Is the word $g(f^\omega(a))$ ultimately periodic ?

Periodicity problem

Let A be a finite alphabet.

Problem (D0L periodicity problem)

If $f : A \to A^*$ is a prolongable morphism on $a \in A$,
is the infinite word $f^\omega(a)$ ultimately periodic ?

It is decidable.
[Harju, Linna, 1986] [Pansiot, 1986]

Baum-Sweet sequence

Let $A = \{a, b, c, d\}$, $B = \{0, 1\}$,

$$f: \begin{array}{l} a \mapsto ab \\ b \mapsto cb \\ c \mapsto bd \\ d \mapsto dd \end{array} \quad \text{and} \quad g: \begin{array}{l} a \mapsto 1 \\ b \mapsto 1 \\ c \mapsto 0 \\ d \mapsto 0 \end{array}$$

$$\begin{array}{l} a \\ \xrightarrow{f} ab \\ \xrightarrow{f} abcb \\ \xrightarrow{f} abcbbdcb \\ \xrightarrow{f} \vdots \end{array}$$

We have

$$f^{\omega}(a) = abcbbdcbcbddbd \cdots$$

$$\begin{aligned} (x_n)_{n \geq 0} &:= g(f^{\omega}(a)) \\ &= 11011001010010 \ldots \end{aligned}$$

k-automatic sequences

Let $k \geq 2$.

A morphism $f : A \to A^*$ is k-uniform if $|f(\alpha)| = k \; \forall \alpha \in A$.

Theorem (Cobham, 1972)

An infinite word x is k-automatic iff there exist

- *a k-uniform morphism $f : A \to A^*$ prolongable on $a \in A$,*
- *a coding $g : A \to B$ such that $x = g(f^\omega(a))$.*

A sequence $(x_n)_{n \geq 0}$ is k-automatic if the n-th term x_n is obtained by feeding a DFA with output with the base k representation of n.

Baum-Sweet sequence

$(x_n)_{n\geq 0} = 11011001010010\ldots$

x_n	1	1	0	1	1	0	0	1	0	...
n	0	1	2	3	4	5	6	7	8	...
$\mathrm{rep}_2(n)$	ε	1	10	11	100	101	110	111	1000	$\cdots$

$x_n = 1 \Leftrightarrow$
No block of 0 of odd length appears in $\mathrm{rep}_2(n)$

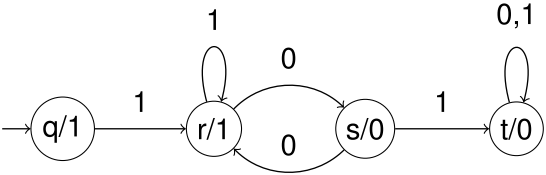
Baum-Sweet sequence

$$(x_n)_{n\geq 0} = 11011001010010\ldots$$

x_n	1	1	0	1	1	0	0	1	0	...
n	0	1	2	3	4	5	6	7	8	...
$\mathrm{rep}_2(n)$	ε	1	10	11	100	101	110	111	1000	...

$x_n = 1 \Leftrightarrow$

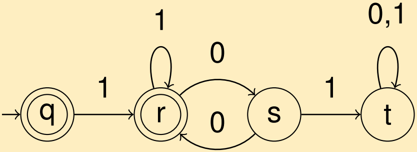
No block of 0 of odd length appears in $\mathrm{rep}_2(n)$



$\rightsquigarrow$ $(x_n)_{n\geq 0}$ is 2-automatic.

Example (Baum-Sweet sequence)

$S_1 = \{n \geq 0 \mid x_n = 1\}$ is 2-recognizable,
i.e., $\mathrm{rep}_2(S_1) = \{\mathrm{rep}_2(n) \mid n \in S_1\}$ is accepted by a DFA.



Conversely,

$$X \subseteq \mathbb{N} \ k\text{-recognizable} \Rightarrow 1_X \ k\text{-automatic}.$$

Problem

Let $g : A \to B$ be a coding and

$f : A \to A^*$ be a k-uniform morphism prolongable on $a \in A^*$.

Is the word $g(f^\omega(a))$ ultimately periodic ?

A set $X \subseteq \mathbb{N}$ is ultimately periodic if 1_X is ultimately periodic.

Equivalent problem

Given a DFA that accepts the base k representation of $X \subseteq \mathbb{N}$, is the set X ultimately periodic ?

Integer base

Problem

Given a DFA that accepts the base k representation of $X \subseteq \mathbb{N}$, is the set X ultimately periodic ?

It is decidable. [Honkala, 1986]

Integer base

Problem

Given a DFA that accepts the base k representation of $X \subseteq \mathbb{N}$, is the set X ultimately periodic ?

It is decidable. [Honkala, 1986]

Remark

Allouche, Rampersad, Shallit	2009
Leroux	2005
Muchnik	1991

Integer base

<u>Idea</u> :

If $X \subseteq \mathbb{N}$ is ultimately periodic,
then the state complexity of the DFA $\nearrow$ with the period and preperiod of X.

<u>Decision method</u> :

Input : $X \subseteq \mathbb{N}$ given by a DFA accepting $0^* rep_b(X)$.

If X is ultimately periodic,
we have an upper bound on its period and its preperiod.

$\rightsquigarrow$ a finite number of pairs (period, preperiod) to test.

General case

Abstract numeration system $S = (L, \Sigma, <)$ where

- L is infinite recognizable language,
- $(\Sigma, <)$ is a totally ordered alphabet.

The representation of an integer n is

$$\mathrm{rep}_S(n) := \text{ the } (n+1)\text{-th word of L.}$$

Example

Let $S = (L, \{a, b\}, a < b)$ with $L = \{\varepsilon\} \cup \{a, ab\}^*$.

$\mathrm{rep}_S(\mathbb{N})$	ε	a	aa	ab	aaa	aab	aba	$aaaa$	$\dots$
$\mathbb{N}$	0	1	2	3	4	5	6	7	$\dots$

Back to the problem

General problem (HD0L periodicity problem)

Let

- $g : A \to B$ be a coding ,
- $f : A \to A^*$ be a morphism prolongable on $a \in A^*$.

Is the word $g(f^\omega(a))$ ultimately periodic ?

Theorem (Maes, Rigo, 2002)

An infinite word x is S-automatic iff there exist

- *a morphism $f : A \to A^*$ prolongable on $a \in A$,*
- *a coding $g : A \to B$ such that $x = g(f^\omega(a))$.*

Problem (equivalent to the "HD0L periodicity problem")

Let

- S be an abstract numeration system,
- $X \subseteq \mathbb{N}$ be a set such that $\mathrm{rep}_S(X)$ is recognizable.

Is the set X ultimately periodic ?

It is decidable for a class of abstract numeration systems.
[Bell, Charlier, Fraenkel, Rigo, 2008]

Positional numeration system

A positional numeration system $U = (U_i)_{i \geq 0}$ is

- a strictly increasing sequence U of integers such that
- $\{U_{i+1}/U_i \mid i \geq 0\}$ is bounded,
- $U_0 = 1$.

Remark

Particular case : integer base

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

Fibonacci numeration system

Let $F = (F_i)_{i \geq 0} := (1, 2, 3, 5, 8, 13, 21, 34, \ldots)$ be given by

$$F_0 = 1, F_1 = 2 \text{ and } F_{i+2} = F_{i+1} + F_i \text{ for all } i \geq 0.$$

13	8	5	3	2	1	
					ε	0
					1	1
				1	0	2
			1	0	0	3
			1	0	1	4
						$\vdots$
1	0	0	1	0	1	17

$\mathrm{rep}_F(17) = 100101$

Fibonacci numeration system

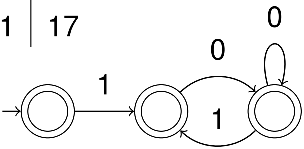
Let $F = (F_i)_{i\geq 0} := (1, 2, 3, 5, 8, 13, 21, 34, \ldots)$ be given by

$$F_0 = 1, F_1 = 2 \text{ and } F_{i+2} = F_{i+1} + F_i \text{ for all } i \geq 0.$$

13	8	5	3	2	1	
					ε	0
					1	1
				1	0	2
			1	0	0	3
			1	0	1	4
						$\vdots$
1	0	0	1	0	1	17

$$\mathrm{rep}_F(17) = 100101$$

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



Positional numeration system

Problem

Let

- U be a positional numeration system,
- $X \subseteq \mathbb{N}$ be a set such that $\mathrm{rep}_U(X)$ is recognizable.

Is the set X ultimately periodic ?

It is decidable for a class of positional numeration systems.
[Bell, Charlier, Fraenkel, Rigo, 2008]

Remark

The decision procedure of Bell *et al.* can not be applied to the integer base systems.

Let L be a language over the finite alphabet A.

Context of a word $u \in A^*$ with respect to L :

$$C_L(u) = \{(x, y) \in A^* \times A^* | xuy \in L\}$$

Myhill congruence for L : $\forall u, v \in A^*$,

$$u \leftrightarrow_L v \Leftrightarrow C_L(u) = C_L(v)$$

Let L be a language over the finite alphabet A.

Context of a word $u \in A^*$ with respect to L :

$$C_L(u) = \{(x,y) \in A^* \times A^* | xuy \in L\}$$

Myhill congruence for L : $\forall u, v \in A^*$,

$$u \leftrightarrow_L v \Leftrightarrow C_L(u) = C_L(v)$$

Example

Let $A = \{a, b\}$ and $L = a^* b^* = \{a^n b^m \,|\, n, m \in \mathbb{N}\}$.

$$\begin{aligned} C_L(ab) &= \{\, (a^i, b^j) \,|\, i, j \in \mathbb{N}\} \\ C_L(ba) &= \emptyset \\ C_L(a) &= \{(a^i, a^j b^\ell) | i, j, \ell \in \mathbb{N}\} \end{aligned}$$

Let $[u]$ denote the class of $u \in A^*$ in $A^*/_{\leftrightarrow_L}$.

The product is defined by

$$[u] \circ [v] = [w] \text{ if } [u] \cdot [v] \subseteq [w].$$

In particular, $[u] \circ [v] = [uv]$.

Syntactic monoid of L : $(A^*/_{\leftrightarrow_L}, \circ)$

Theorem

L is recognizable $\Leftrightarrow$ $A^*/_{\leftrightarrow_L}$ is finite

Syntactic complexity of L : $\#(A^*/_{\leftrightarrow_L})$

Back to the problem

Problem

Given a DFA that accepts the representation of $X \subseteq \mathbb{N}$, is the set X ultimately periodic ?

If $X \subseteq \mathbb{N}$ is periodic of period m,

then the representation of X in a reasonable numeration system gives a language $L \subseteq A^*$ recognizable by a DFA.

Question : Does $\#(A^*/_{\leftrightarrow_L})$ grow with the period m of X ?

Integer base

Theorem (Rigo, V., 2011)

Let $m, b \geq 2$ be integers such that $(m, b) = 1$. If $X \subseteq \mathbb{N}$ is periodic of period m, then

$$\#(A^*/_{\leftrightarrow_{0^*\mathrm{rep}_b(X)}}) = m \cdot \mathrm{ord}_m(b).$$

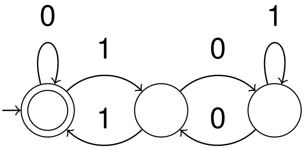
Notation : $\mathrm{ord}_m(b) = \min\{j \in \mathbb{N} \setminus \{0\} \mid b^j \equiv 1 \pmod{m}\}$.

Idea : Show for all $u, v \in A^*$,

$$u \leftrightarrow_{0^*rep_b(X)^*} v \Leftrightarrow \begin{cases} \mathrm{val}_b(u) \equiv \mathrm{val}_b(v) & \pmod{m} \\ |u| \equiv |v| & \pmod{\mathrm{ord}_m(b)} \end{cases}.$$

Example : $X = 3\,\mathbb{N} = \{3n \mid n \in \mathbb{N}\}$

- $b = 2$
- $m = 3$
- $\mathrm{ord}_3(2) = 2$



Multiplication table of the syntactic monoid of $0^*\mathrm{rep}_2(X)$:

	ε	0	1	01	10	101
ε	ε	0	1	01	10	101
0	0	ε	01	1	101	10
1	1	10	ε	101	0	01
01	01	101	0	10	ε	1
10	10	1	101	ε	01	0
101	101	01	10	0	1	ε

Integer base

Lower bounds on the syntactic complexity can be obtained for a period m and a base b such that :

- $m = q$ with $\gcd(q, b) = 1$,
- $m = b^n$ with $n \geq 1$,
- $m = b^n q$ with $q \geq 2$, $\gcd(q, b) = 1$ and $n \geq 1$,
- $m = db^n q$ with $q \geq 2$, $\gcd(q, b) = 1$, $\gcd(d, b) \geq 1$ and $n \geq 0$.

Integer base

Proposition (Rigo, V., 2011)

If b is prime and $X \subseteq \mathbb{N}$ is ultimately periodic of period $m = qb^n$ with $q \geq 2$, $\gcd(q, b) = 1$ and $n \geq 0$, then

$$\#(A^*/_{\leftrightarrow_{0^*\mathrm{rep}_b(X)}}) \geq (n+1)q.$$

In the proof, we use a result of Perles, Rabin, Shamir (1963) on

n-definite languages a. k. a. suffix testable languages.

[Pin, 1997]

A language $L \in A^*$ is a *n-definite* if

- $\forall\, u, v \in A^*$ such that $u = u'x$, $v = v'x$ with $|x| = n$ and $u', v' \in A^*$,
$$u \in L \Leftrightarrow v \in L$$
- $\exists\, u, v \in A^*$ such that $u = u'x$, $v = v'x$ with $|x| = n - 1$ and $u', v' \in A^* \setminus \{\varepsilon\}$,
$$u \in L \text{ and } v \notin L.$$

Example

Let $X = 5 + 8\,\mathbb{N}$ and $L = 0^* rep_2(X)$.
L is 3-definite because

$$L = \{0, 1\}^*\{101\}$$

Goal : Deal with a larger class of numeration systems using the syntactic monoid.

Goal : Deal with a larger class of numeration systems using the syntactic monoid.

Conjecture (Fibonacci numeration system)

$$F_0 = 1, F_1 = 2 \text{ and } F_{i+2} = F_{i+1} + F_i \text{ for all } i \geq 0.$$

If $X = m\mathbb{N} = \{m \cdot n \mid n \in \mathbb{N}\}$, *then*

$$\#(A^*/_{\leftrightarrow_{0^*\mathrm{rep}_F(X)}}) = 4 \cdot m^2 \cdot P_F(m) + 2$$

where $P_F(m)$ *is the period of* $(F_i \bmod m)_{i \geq 0}$.

Thank you.