A decision problem for ultimately periodic sets in non-standard numeration systems

Émilie Charlier

Université libre de Bruxelles

Decidability workshop, Amiens, 3 juillet 2012

Background

Let's start with classical k-ary numeration system, $k \geq 2$:

$$n = \sum_{i=0}^{\ell} d_i\, k^i,\ d_\ell \neq 0, \quad \mathrm{rep}_k(n) = d_\ell \cdots d_0 \in \{0, \ldots, k-1\}^*$$

A set $X \subseteq \mathbb{N}$ is *k-recognizable*, if the language

$$\mathrm{rep}_k(X) = \{\mathrm{rep}_k(x) \colon x \in X\}$$

is regular, i.e. accepted by a finite automaton.

Background

Examples of k-recognizable sets

- In base 2, the even integers: $\mathrm{rep}_2(2\mathbb{N}) = 1\{0,1\}^*0 \cup \{\varepsilon\}$
- In base 2, the powers of 2: $\mathrm{rep}_2(\{2^i | i \in \mathbb{N}\}) = 10^*$
- In base 2, the *Thue-Morse set*:
  $\{n \in \mathbb{N} \colon \mathrm{rep}_2(n)$ contains an even numbers of 1s$\}$
- Given a k-automatic sequence $(x_n)_{n\geq 0}$ over an alphabet Σ, then, for all $\sigma \in \Sigma$, the set $\{i \in \mathbb{N} \colon x_i = \sigma\}$ is k-recognizable.

Background

Divisibility criteria

If $X \subseteq \mathbb{N}$ is ultimately periodic, then X is k-recognizable $\forall k \geq 2$.

$$X = (3\mathbb{N} + 1) \cup (2\mathbb{N} + 2) \cup \{3\}, \text{ Preperiod} = 4, \text{ Period} = 6$$

$\chi_X =$ 

Two integers $k,\ \ell \geq 2$ are *multiplicatively independent* if $k^m = \ell^n \Rightarrow m = n = 0$.

Theorem (Cobham 1969)

Let $k, \ell \geq 2$ be multiplicatively independent integers.
If $X \subseteq \mathbb{N}$ is k- and ℓ-recognizable, then X is ultimately periodic.

Start for this work

Theorem (J. Honkala 1986)

It is decidable if a k-recognizable set is ultimately periodic.

Sketch of Honkala's decision procedure:

- The input is a DFA $\mathcal{A}_X$ accepting $\mathrm{rep}_k(X)$.
- The number of states of $\mathcal{A}_X$ produces an upper bound on the possible (minimal) preperiod and period for X.
- Consequently, there are finitely many candidates to check.
- For each pair (a, p) of candidates, produce a DFA for all possible corresponding ultimately periodic sets and compare it with $\mathcal{A}_X$.

Non standard numeration systems

- A *numeration system* (NS) is an increasing sequence of integers $U = (U_n)_{n \geq 0}$ such that
  - $U_0 = 1$ and
  - $C_U := \sup_{n \geq 0} \lceil U_{n+1}/U_n \rceil < +\infty$.
- U is *linear* if it satisfies a linear recurrence relation over $\mathbb{Z}$.
- Let $n \in \mathbb{N}$. A word $w = w_{\ell-1} \cdots w_0$ over $\mathbb{N}$ represents n if

$$\sum_{i=0}^{\ell-1} w_i \, U_i = n.$$

- In this case, we write $\mathrm{val}_U(w) = n$.

Greedy representations

- A representation $w = w_{\ell-1} \cdots w_0$ of an integer is *greedy* if

$$\forall j, \sum_{i=0}^{j-1} w_i U_i < U_j.$$

- In that case, $w \in \{0, 1, \ldots, C_U - 1\}^*$.
- $\mathrm{rep}_U(n)$ is the greedy representation of n with $w_{\ell-1} \neq 0$.
- $X \subseteq \mathbb{N}$ is U-recognizable $\stackrel{\Delta}{\Leftrightarrow}$ $\mathrm{rep}_U(X)$ is accepted by a finite automaton.
- $\mathrm{rep}_U(\mathbb{N})$ is the numeration language.

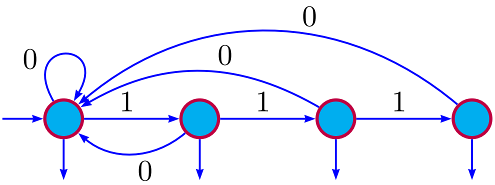
Example (Zeckendorf system)

It is based on the sequence $F = (F_i)_{i\geq 0} = (1, 2, 3, 5, 8, 13, \ldots)$ defined by $F_0 = 1$, $F_1 = 2$ and $F_{i+2} = F_{i+1} + F_i$ for all $i \geq 0$.

1	1	8	10000	15	100010
2	10	9	10001	16	100100
3	100	10	10010	17	100101
4	101	11	10100	18	101000
5	1000	12	10101	19	101001
6	1001	13	100000	20	101010
7	1010	14	100001	21	1000000

The “pattern” 11 is forbidden, $A_F = \{0, 1\}$.

The ℓ-bonacci numeration system



- $U_{n+\ell} = U_{n+\ell-1} + U_{n+\ell-2} + \cdots + U_n$
- $U_i = 2^i$, $i \in \{0, \dots, \ell - 1\}$
- $\mathcal{A}_U$ accepts all words that do not contain 1^ℓ.

A decision problem

Proposition

Let $U = (U_i)_{i \geq 0}$ be a NS s.t. $\mathbb{N}$ is U-recognizable. Any ultimately periodic $X \subseteq \mathbb{N}$ is U-recognizable and a DFA accepting $\mathrm{rep}_U(X)$ can be obtained effectively.

NB: If $\mathbb{N}$ is U-recognizable, then U is linear.

Periodicity problem: Given U s.t. $\mathbb{N}$ is U-recognizable and a U-recognizable set $X \subseteq \mathbb{N}$. Is it decidable if X is ultimately periodic ?

First part: an upper bound on the period

"Pseudo-result"

Let X be ultimately periodic with period p_X.
Any DFA accepting $\mathrm{rep}_U(X)$ has at least $f(p_X)$ states,
where f is increasing.

"Pseudo-corollary"

Let $X \subseteq \mathbb{N}$ be a U-recognizable set of integers s.t. $\mathrm{rep}_U(X)$ is accepted by a d-state DFA.
If X is ultimately periodic with period p_X, then

$$\boxed{f(p_X) \leq d} \quad \text{with} \begin{cases} d \text{ fixed} \\ f \text{ increasing.} \end{cases}$$

$\Rightarrow$ The number of candidates for the period is bounded from above.

A technical hypothesis :

$$\lim_{i \to +\infty} U_{i+1} - U_i = +\infty. \qquad (1)$$

Most systems are built on an exponential sequence $(U_i)_{i \geq 0}$.

Lemma

Let $U = (U_i)_{i \geq 0}$ be a NS satisfying (1)*. If w is a greedy U-representation, then so is $10^r w$ for all r large enough.*

Let $N_U(m) \in \{1, \ldots, m\}$ denotes the number of values that are taken infinitely often by the sequence $(U_i \bmod m)_{i \geq 0}$.

Example (Zeckendorf system)

$(F_i \bmod 4) = (1, 2, 3, 1, 0, 1, 1, 2, 3, \ldots)$, so $N_F(4) = 4$.

$(F_i \bmod 11) = (1, 2, 3, 5, 8, 2, 10, 1, 0, 1, 1, 2, 3, \ldots)$, so $N_F(11) = 7$.

If $U = (U_i)_{i \geq 0}$ is a linear system of order k, then, for all $m \geq 2$, we have

$$\sqrt[k]{\pi_U(m)} \leq N_U(m) \leq \pi_U(m),$$

where $\pi_U(m)$ denotes the minimal period of $(U_i \bmod m)_{i \geq 0}$.

Theorem (C-Rigo 2008)

Let U be a NS satisfying (1). If $X \subseteq \mathbb{N}$ is an ultimately periodic U-recognizable set of period p_X, then any DFA accepting $\mathrm{rep}_U(X)$ has at least $N_U(p_X)$ states.

Corollary

Let U be a NS satisfying (1). Assume that

$$\lim_{m \to +\infty} N_U(m) = +\infty.$$

Then the period of an ultimately periodic set $X \subseteq \mathbb{N}$ s.t. $\mathrm{rep}_U(X)$ is accepted by a d-state DFA is bounded by the smallest integer M s.t. $N_U(m) > d$ for all $m \geq M$, which is effectively computable.

Proposition

If $U = (U_i)_{i \geq 0}$ *satisfies a recurrence relation of the kind*

$$U_{i+k} = a_1 U_{i+k-1} + \cdots + a_k U_i, \tag{2}$$

with $a_k = \pm 1$*, then* $\lim_{m \to +\infty} N_U(m) = +\infty$*.*

Proposition

Let $U = (U_i)_{i \geq 0}$ *be an increasing sequence satisfying (2). The following assertions are equivalent:*

- $\lim_{m \to +\infty} N_U(m) = +\infty$;
- *for all prime divisors* p *of* a_k, $\lim_{v \to +\infty} N_U(p^v) = +\infty$.

A characterization

Let $Q_U(x)$ denote the characteristic polynomial of the shortest recurrence relation satisfied by U; and let $P_U(x) = x^k Q_U(\frac{1}{x})$, where $k = \deg(Q_U(x))$.

Theorem (Bell-C-Fraenkel-Rigo 2009)

We have $N_U(p^v) \to +\infty$ as $v \to +\infty$ if and only if

$$P_U(x) = A(x)B(x)$$

with $A(x), B(x) \in \mathbb{Z}[x]$ such that:

- $B(x) \equiv 1 \pmod{p\,\mathbb{Z}[x]}$;
- $A(x)$ *has no repeated roots and all its roots are roots of unity.*

Logical approach

Theorem (Muchnik 1991)

The ultimate periodicity problem is decidable for all NS with a regular numeration language, provided that addition is recognizable.

Example ($U_{i+4} = 3U_{i+3} + 2U_{i+2} + 3U_i$ for all $i \geq 0$, $(U_0, U_1, U_2, U_3) = (1, 2, 3, 4)$)

Addition is not computable by a finite automaton (due to Frougny).
Nevertheless, $N_U(3^v) \to +\infty$ as $v \to +\infty$ because

$$P_U = 1 - 3x - 2x^2 - 3x^4$$

cannot be factorized as $A \cdot B$ with two factors satisfying the hypotheses of the characterization mentioned above.

One of the main arguments for the decidability

Theorem (C-Rigo 2008)

Let U be a NS satisfying (1) and $X \subseteq \mathbb{N}$ be an ultimately periodic U-recognizable set of period p_X. If 1 occurs infinitely many times in $(U_i \bmod p_X)_{i\geq 0}$ then any DFA accepting $\mathrm{rep}_U(X)$ has at least p_X states.

Idea of the proof with the Zeckendorf system

Theorem (Zeckendorf system)

Let $X \subseteq \mathbb{N}$ be ultimately periodic with period p_X (and preperiod a_X). Any DFA accepting $\mathrm{rep}_F(X)$ has at least p_X states.

- $w^{-1}L = \{u \colon wu \in L\} \leftrightarrow$ states of minimal automaton of L
- $(F_i \bmod p_X)_{i \geq 0}$ is *purely periodic*.
- If $i, j \geq a_X$ and $i \not\equiv j \pmod{p_X}$ then there exists $t < p_X$ s.t. either $i + t \in X$ and $j + t \notin X$, or $i + t \notin X$ and $j + t \in X$.
- $\exists n_1, \ldots, n_{p_X}$, $\forall t$, $0 \leq t < p_X$, the words

$$10^{n_{p_X}} \cdots 10^{n_2} 10^{n_1} 0^{|\mathrm{rep}_F(p_X - 1)| - |\mathrm{rep}_F(t)|} \mathrm{rep}_F(t)$$

are greedy F-representations.

Idea of the proof with the Zeckendorf system

- Moreover $n_1, \dots, n_{p_X}$ can be chosen s.t. $\forall j$, $1 \leq j \leq p_X$,

$$\mathrm{val}_F(10^{n_j} \cdots 10^{n_1 + |\mathrm{rep}_F(p_X - 1)|}) \equiv j \pmod{p_X}$$

  and $\mathrm{val}_F(10^{n_1 + |\mathrm{rep}_F(p_X - 1)|}) \geq a_X$.

- For $i, j \in \{1, \dots, p_X\}$, $i \neq j$, the words

$$10^{n_i} \cdots 10^{n_1} \text{ and } 10^{n_j} \cdots 10^{n_1}$$

  will generate different states in the minimal automaton of $\mathrm{rep}_F(X)$. This can be shown by concatenating some word of length $|\mathrm{rep}_F(p_X - 1)|$.

$w^{-1}L = \{u \colon wu \in L\} \leftrightarrow$ states of minimal automaton of L

$X = (11\mathbb{N} + 3) \cup \{2\}$, $a_X = 3$, $p_X = 11$, $|\operatorname{rep}_F(10)| = 5$

Working in $(F_i \bmod 11)_{i \geq 0}$:

$\cdots$ 2 1	1 0 1 10 2 8 5 3 2 1	1 0 1 10 2 8 5 3 2 1	
	1	0 0 0 0 0 0 0 0 0 0	1
1	0 0 0 0 0 0 0 0 0 1	0 0 0 0 0 0 0 0 0 0	2
	1	0 0 0 0 0 0 0 0 1 0	1+2 $\in X$
1	0 0 0 0 0 0 0 0 0 1	0 0 0 0 0 0 0 0 1 0	2+2 $\notin X$

$$\Rightarrow (10^5)^{-1} \operatorname{rep}_F(X) \neq (10^9 10^5)^{-1} \operatorname{rep}_F(X)$$

Second part: an upper bound on the preperiod

For a sequence $U = (U_i)_{i\geq 0}$ of integers, if $(U_i \bmod m)_{i\geq 0}$ is ultimately periodic, we denote its (minimal) preperiod by $\iota_U(m)$.

Theorem (C-Rigo 2008)

Let $U = (U_i)_{i\geq 0}$ be a linear numeration system. Let $X \subseteq \mathbb{N}$ be ultimately periodic with period p_X and preperiod a_X. Then any DFA accepting $\mathrm{rep}_U(X)$ has at least

$$|\mathrm{rep}_U(a_X - 1)| - \iota_U(p_X) \text{ states.}$$

If p_X is bounded, then the number of states grows as a_X grows.

A Decision Procedure

Theorem (C-Rigo 2008)

It is decidable if a U-recognizable set is ultimately periodic for numeration systems $U = (U_i)_{i \geq 0}$ s.t.

- $\mathbb{N}$ *is U-recognizable;*
- $\lim_{i \to +\infty} U_{i+1} - U_i = +\infty$;
- $\lim_{m \to +\infty} N_U(m) = +\infty$.

Further work

Remark

Whenever $\gcd(a_1, \ldots, a_k) = g \geq 2$, we have $U_i \equiv 0 \pmod{g^n}$ for all $n \geq 1$ and for all i large enough; hence $N_U(m) \not\to +\infty$.

Examples

- Integer bases: $U_{n+1} = k\, U_n$
- $U_{n+2} = 2U_{n+1} + 2U_n$

$$a, b, 2(a+b), 2(2a+3b), 4(3a+4b), 4(8a+11b) \ldots$$

Some related references

Learn more about linear recurrent sequences mod m ...

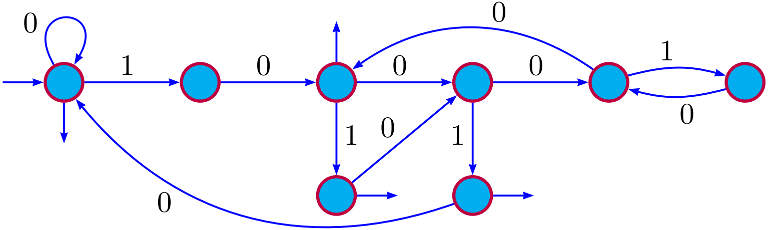
- H.T. Engstrom, On sequences defined by linear recurrence relations, *Trans. Amer. Math. Soc.* **33** (1931).
- M. Ward, The characteristic number of a sequence of integers satisfying a linear recursion relation, *Trans. Amer. Math. Soc.* **35** (1933).
- M. Hall, An isomorphism between linear recurring sequences and algebraic rings, *Trans. Amer. Math. Soc.* **44** (1938).
- G. Rauzy, Relations de récurrence modulo m, Séminaire Delange-Pisot, 1963/1964.

Transition to state complexity

Main ideas for an automata-resolution of the periodicity problem:

- If $X \subseteq \mathbb{N}$ is ultimately periodic, then the state complexity of the associated minimal DFA should grow with the period and preperiod of X.
- Analyse the inner structure of DFAs accepting the U-representations of $m\,\mathbb{N} + r$.

F-representations of even numbers



13	8	5	3	2	1	
				1	0	2
			1	0	1	4
		1	0	0	1	6
	1	0	0	0	0	8
	1	0	0	1	0	10
	1	0	1	0	1	12
						⋮