A DECISION PROBLEM FOR ULTIMATELY PERIODIC SETS IN NON-STANDARD NUMERATION SYSTEMS

JASON BELL, EMILIE CHARLIER, AVIEZRI S. FRAENKEL, AND MICHEL RIGO

ABSTRACT. Consider a non-standard numeration system like the one built over the Fibonacci sequence where nonnegative integers are represented by words over $\{0, 1\}$ without two consecutive 1. Given a set X of integers such that the language of their greedy representations in this system is accepted by a finite automaton, we consider the problem of deciding whether or not X is a finite union of arithmetic progressions. We obtain a decision procedure for this problem, under some hypothesis about the considered numeration system. In a second part, we obtain an analogous decision result for a particular class of abstract numeration systems built on an infinite regular language.

1. INTRODUCTION

Definition 1. A *positional numeration system* is given by a (strictly) increasing sequence $U = (U_i)_{i \geq 0}$ of integers such that $U_0 = 1$ and $C_U := \sup_{i \geq 0} \lceil U_{i+1}/U_i \rceil$ is finite. Let $A_U = \{0, \ldots, C_U - 1\}$. The *greedy U-representation* of a positive integer n is the unique finite word $\mathrm{rep}_U(n) = w_\ell \cdots w_0$ over A_U satisfying

$$n = \sum_{i=0}^{\ell} w_i \, U_i, \ w_\ell \neq 0 \text{ and } \sum_{i=0}^{t} w_i \, U_i < U_{t+1}, \ \forall t = 0, \ldots, \ell.$$

We set $\mathrm{rep}_U(0)$ to be the empty word ε. A set $X \subseteq \mathbb{N}$ of integers is *U-recognizable* if the language $\mathrm{rep}_U(X)$ over A_U is regular (i.e., accepted by a deterministic finite automaton, DFA). If $x = x_\ell \cdots x_0$ is a word over a finite alphabet of integers, then the *U-numerical value* of x is

$$\mathrm{val}_U(x) = \sum_{i=0}^{\ell} x_i \, U_i.$$

Remark 2. Let x, y be two words over A_U. As a consequence of the greediness of the representation, if xy is a greedy U-representation and if the leftmost letter of y is not 0, then y is also a greedy U-representation. Notice that for $m, n \in \mathbb{N}$, we have $m < n$ if and only if $\mathrm{rep}_U(m) <_{gen} \mathrm{rep}_U(n)$ where $<_{gen}$ is the genealogical ordering over A_U^*: words are ordered by increasing length and for words of same length, one uses the lexicographical ordering induced by the natural ordering of the digits in the alphabet A_U. Recall that for two words $x, y \in A_U^*$ of same length, x is lexicographically smaller than y if there exist $w, x', y' \in A_U^*$ and $a, b \in A_U$ such that $x = wax'$, $y = wby'$ and $a < b$.

For a positional numeration system U, it is natural to expect that $\mathbb{N}$ is U-recognizable. A necessary condition is that the sequence U satisfies a linear recurrence relation [39].

Definition 3. A positional numeration system $U = (U_i)_{i \geq 0}$ is said to be *linear*, if the sequence U satisfies a homogenous linear recurrence relation with integer

coefficients. Otherwise stated, there exist $k \geq 1$ and constant coefficients $a_1, \ldots, a_k$ such that for all $i \geq 0$, we have

$$(1) \qquad U_{i+k} = a_1 U_{i+k-1} + \cdots + a_k U_i, \quad \text{with } a_1, \ldots, a_k \in \mathbb{Z}, \ a_k \neq 0.$$

We say that k is the *order* of the recurrence relation.

Example 4. Consider the sequence defined by $F_0 = 1$, $F_1 = 2$ and for all $i \geq 0$, $F_{i+2} = F_{i+1} + F_i$. The *Fibonacci (linear numeration) system* is given by $F = (F_i)_{i \geq 0} = (1, 2, 3, 5, 8, 13, \ldots)$. For instance, $\mathrm{rep}_F(15) = 100010$ and $\mathrm{val}_F(101001) = 13 + 5 + 1 = 19$.

In this paper, we mainly address the following decidability question and its extension to abstract numeration systems.

Problem 1. Given a linear numeration system U and a set $X \subseteq \mathbb{N}$ such that $\mathrm{rep}_U(\mathbb{N})$ and $\mathrm{rep}_U(X)$ are both recognized by (deterministic) finite automata. Is it decidable whether or not X is ultimately periodic, i.e., whether or not X is a finite union of arithmetic progressions ?

Notice that the regularity of $\mathrm{rep}_U(\mathbb{N})$ ensures that there exists a set $X \subseteq \mathbb{N}$ such that $\mathrm{rep}_U(X)$ is regular, see also Remark 20.

Ultimately periodic sets of integers play a special role. On the one hand such infinite sets are coded by a finite amount of information. On the other hand the celebrated Cobham's theorem asserts that these sets are the only sets that are recognizable in all integer base systems [12]. This is the reason why they are also referred in the literature as *recognizable* sets of integers (the recognizability being in that case independent of the base). Moreover, Cobham's theorem has been extended to various situations and in particular, to numeration systems given by substitutions [13].

If we restrict ourselves to the usual integer base $b \geq 2$ numeration system, referred to in the sequel as *b-ary system*, defined by $U_i = b \, U_{i-1}$ for $i \geq 1$ and $U_0 = 1$, several results are known. J. Honkala showed in [26] that Problem 1 turns out to be decidable. Let us also mention [1], where the number of states of the minimal automaton accepting numbers written in base b and divisible by d is given explicitly. J.-P. Allouche and J. Shallit ask in [2] if one can obtain a polynomial time decision procedure for b-ary systems. Using the logic formalism of the Presburger arithmetic, a positive answer to this question is given by J. Leroux in [30] even when considering subsets of $\mathbb{Z}^d$, $d \geq 1$. In dimension one, ultimately periodic sets are exactly the sets definable in the Presburger arithmetic $\langle \mathbb{N}, + \rangle$.

Note that A. Muchnik has shown that Problem 1 turns out to be decidable for any linear numeration system U for which both $\mathrm{rep}_U(\mathbb{N})$ and addition are recognizable by automata [32]. But it is a difficult question to characterize numeration systems U for which addition is *computable by finite automaton*, i.e., for which the language

$$\left\{ \begin{pmatrix} 0^{m - |\mathrm{rep}_U(x)|} \, \mathrm{rep}_U(x) \\ 0^{m - |\mathrm{rep}_U(y)|} \, \mathrm{rep}_U(y) \\ 0^{m - |\mathrm{rep}_U(z)|} \, \mathrm{rep}_U(z) \end{pmatrix} \mid x, y, z \in \mathbb{N}, \ x + y = z, \ m = \max_{t \in \{x,y,z\}} |\mathrm{rep}_U(t)| \right\}$$

where leading zeroes are prefixed for padding the shorter components to obtain three words of the same length, is regular (see for instance [7, 21] in which the characteristic polynomial of the sequence U is the minimal polynomial of a Pisot number). In [20], the sequentiality of the successor function (i.e., the action of adding 1) is studied. If addition is computable by a finite automaton, so is the successor function, but the converse does not hold in general. In particular, some examples of linear numeration systems for which addition is not computable by a finite automaton are given in [20]: for instance, the sequence defined by $U_i = 3U_{i-1} + 2U_{i-2} + 3U_{i-4}$

with integer initial conditions $1 = U_0 < U_1 < U_2 < U_3$. So the decision techniques from [30, 32] cannot be applied to that system. Nevertheless, as we will see in Example 32, our decision procedure can be applied to this system. Notice also that in the extended framework of abstract numeration systems, one can exhibit systems such that multiplication by a constant does not preserve recognizability [9, 29, 35]. Therefore the powerful tools from logic discussed above cannot be applied in that context either.

The question studied in this paper was raised by J. Sakarovitch during the "*Journées de Numération*" in Graz, May 2007. The question was initially asked for a larger class of systems than the one treated here, namely for any abstract numeration systems defined on an infinite regular language [29]. A shorter and partial version of this paper has been presented to the 33rd MFCS symposium in Toruń, August 2008 [8].

Observe that this decision problem for all abstract numeration systems is equivalent to the famous *HD0L periodicity problem*: given a morphism f and a coding g, decide if the infinite word $g(f^\omega(a))$ is ultimately periodic, see [27, 36]. For the restricted case of the *D0L periodicity problem*, where only the morphism f is considered, decision procedures are well-known [24, 33]. Finally, questions connected to the ones addressed here have independently and recently gained interest [3]. In particular, a simple proof of Honkala's original result based on the construction of some automata is given in that paper. As for the logical approach considered by Muchnik and Leroux, the arguments given in [3] rely on the recognizability of addition by automata (which can be done for the classical k-ary system but not necessarily for an arbitrary linear numeration system).

The structure of this paper is the same as [26]. First we give an upper bound on the admissible periods of a U-recognizable set X when it is assumed to be ultimately periodic. Then an upper bound on the admissible preperiods is obtained. These bounds depend essentially on the number of states of the (minimal) automaton recognizing $\mathrm{rep}_U(X)$. Finally, finitely many such periods and preperiods have to be checked. For each of them, we have to build an automaton accepting the corresponding ultimately periodic sets (this implies that $\mathbb{N}$ has to be recognizable).

Though the structure is the same, our arguments and techniques are quite different from [26]. They rely on the study of the quantity $N_U(m)$ defined as the number of residue classes that appear infinitely often in the sequence $(U_i \bmod m)_{i \geq 0}$. Our main result — Theorem 22 — can be stated as follows.

Theorem. Let $U = (U_i)_{i \geq 0}$ be a linear numeration system such that $\mathbb{N}$ is U-recognizable and satisfying $\lim_{i \to +\infty} U_{i+1} - U_i = +\infty$. If $\lim_{m \to +\infty} N_U(m) = +\infty$, then Problem 1 is decidable.

Actually our techniques cannot be applied to b-ary systems (see Remark 25), which is the case treated by Honkala [26], because in that case $N_U(m) \not\to +\infty$ as $m \to +\infty$.

In Section 3, we give a characterization of the linear numeration systems U such that $\lim_{m \to +\infty} N_U(m) = +\infty$. To do so we use p-adic methods leading to a study of the sequence $(U_i \bmod p^v)_{i \geq 0}$ for all $v \geq 1$, where p is a prime dividing a_k.

In the last section, we consider again the same decision problem but restated in the framework of abstract numeration systems [29]. The definition of these systems is given in Section 4. We apply successfully the same kind of techniques to a large class of abstract numeration systems (for instance, an example consisting of two copies of the Fibonacci system is considered). The corresponding decision procedure is given by Theorem 42. As explained above, this result provides a decision procedure for specific instances of the HD0L periodicity problem.

All along the paper, we try whenever it is possible to state results in their most general form, even if later on we have to restrict ourselves to particular cases. For instance, results about the admissible preperiods do not require any particular assumption on the numeration system except linearity.

2. A Decision Procedure for a Class of Linear Numeration Systems

We will often consider positional numeration systems $U = (U_i)_{i \geq 0}$ satisfying the following condition:

$$(2) \qquad \lim_{i \to +\infty} U_{i+1} - U_i = +\infty.$$

Notice that it is a weak requirement. Usually, the sequence U has an exponential growth, $U_i \simeq \beta^i$ for some $\beta > 1$, and therefore (2) is trivially satisfied. (It is for instance the case for the numeration systems considered in Remark 27.)

Lemma 5. *Let $U = (U_i)_{i \geq 0}$ be a positional numeration system satisfying* (2). *Then for all j, there exists L such that for all $\ell \geq L$,*

$$10^{\ell - |\operatorname{rep}_U(t)|} \operatorname{rep}_U(t), \ t = 0, \dots, U_j - 1$$

are greedy U-representations. Otherwise stated, if w is a greedy U-representation, then for r large enough, $10^r w$ is also a greedy U-representation.

Proof. Notice that $\operatorname{rep}_U(U_j - 1)$ is the greatest word of length j in $\operatorname{rep}_U(\mathbb{N})$, since $\operatorname{rep}_U(U_j) = 10^j$. By hypothesis, there exists L such that for all $\ell \geq L$, $U_{\ell+1} - U_\ell > U_j - 1$. Therefore, for all $\ell \geq L$,

$$10^{\ell-j} \operatorname{rep}_U(U_j - 1)$$

is the greedy U-representation of $U_\ell + U_j - 1 < U_{\ell+1}$ and the conclusion follows. $\square$

Example 6. Consider the positional numeration system $U_0 = 1$, $U_1 = 2$, $U_2 = 3$ and $U_{3i+r} = 3^{i+1} + r$ for all $i \geq 1$ and $r \in \{0, 1, 2\}$. This system does not satisfy (2) because $U_{i+1} - U_i = 1$ for infinitely many i's. We have $\operatorname{rep}_U(2) = 10$, but one can notice that for $i \geq 0$, $10^{3i+1}10$ is not a greedy U-representation. Indeed, $\operatorname{val}_U(10^{3i+1}10) = U_{3(i+1)} + 2 = U_{3i+5}$ has 10^{3i+5} as greedy U-representation.

Remark 7. In the above lemma, one cannot exchange the order of the quantifiers about j and L. As an example, consider the sequence $U_i = (i+1)(i+2)/2$ for all $i \geq 0$. This sequence satisfies the linear recurrence relation $U_{i+3} = 3U_{i+2} - 3U_{i+1} + U_i$ and also (2). Observe that, for all $i \geq 1$, $k = U_i - 1$ is the unique value such that $U_i = U_k - U_{k-1}$ because for all $j \geq 1$, $U_j - U_{j-1} = j + 1$. For all $i \geq 1$, 10^i is a greedy U-representation and the greedy U-representations of the form $10^n 10^i$ are exactly those for which $n \geq U_i - i - 1$.

Remark 8. Bertrand numeration systems associated with a real number $\beta > 1$ are defined as follows. Let $A_\beta = \{0, \dots, \lceil \beta \rceil - 1\}$. Any $x \in [0, 1]$ can be written as

$$x = \sum_{i=1}^{+\infty} c_i \, \beta^{-i}, \ \text{with } c_i \in A_\beta$$

and the sequence $(c_i)_{i \geq 1}$ is said to be a *β-representation* of x. The maximal β-representation of x for the lexicographical order is denoted $d_\beta(x)$ and is called the *β-development* of x (for details see [31, Chap. 8]). We say that a β-development $(c_i)_{i \geq 1}$ is *finite* if there exists N such that $c_i = 0$ for all $i \geq N$. If there exists $m \geq 1$ such that $d_\beta(1) = t_1 \cdots t_m$ with $t_m \neq 0$, we set $d_\beta^*(1) := (t_1 \cdots t_{m-1}(t_m - 1))^\omega$, otherwise $d_\beta(1)$ is infinite and we set $d_\beta^*(1) := d_\beta(1)$.

We can now define a positional numeration system $U_\beta = (U_i)_{i \geq 0}$ associated with β (see [4]). If $d_\beta^*(1) = (t_i)_{i \geq 1}$, then

(3) $\qquad\qquad U_0 = 1 \text{ and } \forall i \geq 1, \ U_i = t_1 U_{i-1} + \cdots + t_i U_0 + 1.$

If β is a Parry number (i.e., $d_\beta(1)$ is finite or ultimately periodic) then one can derive from (3) that the sequence U_β satisfies a linear recurrence relation and as a consequence of Bertrand's theorem [4] linking greedy U_β-representations and finite factors occurring in β-developments, the language $\mathrm{rep}_{U_\beta}(\mathbb{N})$ of the greedy U_β-representations is regular. The automaton accepting these representations is well-known [22] and has a special form (all states — except for a sink — are final and from all these states, an edge of label 0 goes back to the initial state). We therefore have the following property which is much stronger than the previous lemma. If x and y are greedy U_β-representations then $x0y$ is also a greedy U_β-representation.

Example 9. The Fibonacci system is the Bertrand system associated with the golden ratio $(1 + \sqrt{5})/2$. Since greedy representations in the Fibonacci system are the words not containing two consecutive 1's [41], then for $x, y \in \mathrm{rep}_F(\mathbb{N})$, we have $x0y \in \mathrm{rep}_F(\mathbb{N})$.

Definition 10. Let $X \subseteq \mathbb{N}$ be a set of integers. The *characteristic word of X* is an infinite word $x_0 x_1 x_2 \cdots$ over $\{0, 1\}$ defined by $x_i = 1$ if and only if $i \in X$.

Consider for now $X \subseteq \mathbb{N}$ to be an ultimately periodic set. The characteristic word of X is therefore an infinite word over $\{0, 1\}$ of the form

$$x_0 x_1 x_2 \cdots = uv^\omega$$

where u and v are chosen of minimal length. We say that the length $|u|$ of u (resp. the length $|v|$ of v) is the *preperiod* (resp. *period*) of X. Hence, for all $n \geq |u|$, $n \in X$ if and only if $n + |v| \in X$.

The following lemma is a simple consequence of the minimality of the period chosen to represent an ultimately periodic set.

Lemma 11. *Let $X \subseteq \mathbb{N}$ be an ultimately periodic set of period p_X and preperiod a_X. Let $i, j \geq a_X$. If $i \not\equiv j \mod p_X$ then there exists $t < p_X$ such that either $i + t \in X$ and $j + t \notin X$ or, $i + t \notin X$ and $j + t \in X$.*

We assume that the reader is familiar with automata theory (see for instance [37]) but let us recall some classical results. Let $L \subseteq \Sigma^*$ be a language over a finite alphabet Σ and x be a finite word over Σ. We set

$$x^{-1} L = \{z \in \Sigma^* \mid xz \in L\}.$$

We can now define the Myhill-Nerode congruence. Let $x, y \in \Sigma^*$. We have $x \sim_L y$ if and only if $x^{-1} L = y^{-1} L$. Moreover L is regular if and only if $\sim_L$ has a finite index being the number of states of the minimal automaton of L.

Definition 12. For a sequence $(U_i)_{i \geq 0}$ of integers, $N_U(m) \in \{1, \ldots, m\}$ denotes the number of values that are taken infinitely often by the sequence $(U_i \mod m)_{i \geq 0}$.

Proposition 13. *Let $U = (U_i)_{i \geq 0}$ be a positional numeration system satisfying (2). If $X \subseteq \mathbb{N}$ is an ultimately periodic U-recognizable set of period p_X, then any deterministic finite automaton accepting $\mathrm{rep}_U(X)$ has at least $N_U(p_X)$ states.*

Proof. Let a_X be the preperiod of X. By Lemma 5, there exists L such that for any $h \geq L$, the words

$$10^{h - |\mathrm{rep}_U(t)|} \mathrm{rep}_U(t), \ t = 0, \ldots, p_X - 1$$

are greedy U-representations. The sequence $(U_i \mod p_X)_{i \geq 0}$ takes infinitely often $N := N_U(p_X)$ different values. Let $h_1, \ldots, h_N \geq L$ be such that

$$i \neq j \Rightarrow U_{h_i} \not\equiv U_{h_j} \mod p_X$$

and $h_1, \ldots, h_N$ can be chosen such that $U_{h_i} > a_X$ for all $i \in \{1, \ldots, N\}$.

By Lemma 11, for all $i, j \in \{1, \ldots, N\}$ such that $i \neq j$, there exists $t_{i,j} < p_X$ such that either $U_{h_i} + t_{i,j} \in X$ and $U_{h_j} + t_{i,j} \notin X$, or $U_{h_i} + t_{i,j} \notin X$ and $U_{h_j} + t_{i,j} \in X$. Therefore,

$$w_{i,j} = 0^{|\operatorname{rep}_U(p_X - 1)| - |\operatorname{rep}_U(t_{i,j})|} \operatorname{rep}_U(t_{i,j})$$

is a word such that either

$$10^{h_i - |\operatorname{rep}_U(p_X - 1)|} w_{i,j} \in \operatorname{rep}_U(X) \text{ and } 10^{h_j - |\operatorname{rep}_U(p_X - 1)|} w_{i,j} \notin \operatorname{rep}_U(X),$$

or

$$10^{h_i - |\operatorname{rep}_U(p_X - 1)|} w_{i,j} \notin \operatorname{rep}_U(X) \text{ and } 10^{h_j - |\operatorname{rep}_U(p_X - 1)|} w_{i,j} \in \operatorname{rep}_U(X).$$

Therefore the words $10^{h_1 - |\operatorname{rep}_U(p_X - 1)|}, \ldots, 10^{h_N - |\operatorname{rep}_U(p_X - 1)|}$ are pairwise nonequivalent for the relation $\sim_{\operatorname{rep}_U(X)}$ and the minimal automaton of $\operatorname{rep}_U(X)$ has at least $N = N_U(p_X)$ states. $\square$

The previous proposition has an immediate consequence for getting a bound on the period of a periodic set accepted by a given DFA.

Corollary 14. *Let $U = (U_i)_{i \geq 0}$ be a positional numeration system satisfying (2). Assume that*

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

Then the period of an ultimately periodic set $X \subseteq \mathbb{N}$ such that $\operatorname{rep}_U(X)$ is accepted by a DFA with d states is bounded by the smallest integer s_0 such that for all $m \geq s_0$, $N_U(m) > d$.

A result similar to the previous corollary (in the sense that it permits to give an upper bound on the period) can be stated as follows. This result will be used later on in our decision procedure to compute some explicit estimate.

Proposition 15. *Let $U = (U_i)_{i \geq 0}$ be a positional numeration system satisfying (2) and $X \subseteq \mathbb{N}$ be an ultimately periodic U-recognizable set of period p_X. Let c be a divisor of p_X. If 1 occurs infinitely many times in $(U_i \mod c)_{i \geq 0}$ then any deterministic finite automaton accepting $\operatorname{rep}_U(X)$ has at least c states.*

Proof. Let a_X be the preperiod of X. Applying several times Lemma 5, there exist $n_1, \ldots, n_c$ such that

$$10^{n_c} 10^{n_{c-1}} \cdots 10^{n_1} 0^{|\operatorname{rep}_U(p_X - 1)| - |\operatorname{rep}_U(t)|} \operatorname{rep}_U(t), \ t = 0, \ldots, p_X - 1$$

are greedy U-representations. Moreover, since 1 occurs infinitely many times in the sequence $(U_i \mod c)_{i \geq 0}$, $n_1, \ldots, n_c$ can be chosen such that, for all $j = 1, \ldots, c$,

$$\operatorname{val}_U(10^{n_j} \cdots 10^{n_1 + |\operatorname{rep}_U(p_X - 1)|}) \equiv j \mod c$$

and

$$\operatorname{val}_U(10^{n_1 + |\operatorname{rep}_U(p_X - 1)|}) > a_X.$$

For $i, j \in \{1, \ldots, c\}$, $i \neq j$, by Lemma 11 and since c divides p_X, the words

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

are nonequivalent for $\sim_{\operatorname{rep}_U(X)}$. This can be shown by concatenating some word of the kind $0^{|\operatorname{rep}_U(p_X - 1)| - |\operatorname{rep}_U(t)|} \operatorname{rep}_U(t)$ with $t < p_X$, as in the proof of Proposition 13. This concludes the proof. $\square$

Definition 16. For a sequence $(U_i)_{i \geq 0}$ of integers, if $(U_i \mod m)_{i \geq 0}$ is ultimately periodic, we denote its (minimal) preperiod by $\iota_U(m)$ (we choose the notation ι to allude to the word index which is equally used as preperiod) and its (minimal) period by $\pi_U(m)$.

Remark 17. Observe that for any linear recurrence sequence of order k satisfying (1), we have

$$N_U(m) \leq \pi_U(m) \leq (N_U(m))^k.$$

Therefore, $\lim_{m \to +\infty} N_U(m) = +\infty$ if and only if $\lim_{m \to +\infty} \pi_U(m) = +\infty$. Notice that if $m = p.q$ with $\gcd(p, q) = 1$, then $\pi_U(m) = \text{lcm}\{\pi_U(p), \pi_U(q)\}$.

Now we want to obtain an upper bound on the preperiod of any ultimately periodic U-recognizable set recognized by a given DFA.

Proposition 18. *Let $U = (U_i)_{i \geq 0}$ be a linear numeration system. Let $X \subseteq \mathbb{N}$ be an ultimately periodic U-recognizable set of period p_X and preperiod a_X. Then any deterministic finite automaton accepting $\text{rep}_U(X)$ has at least $|\text{rep}_U(a_X - 1)| - \iota_U(p_X)$ states.*

The arguments of the following proof are similar to the one found in [26].

Proof. W.l.o.g. we can assume that $|\text{rep}_U(a_X - 1)| - \iota_U(p_X) > 0$. The sequence $(U_i \mod p_X)_{i \geq 0}$ is ultimately periodic with preperiod $\iota_U(p_X)$ and period $\pi_U(p_X)$. Proceed by contradiction and assume that $\mathcal{A}$ is a deterministic finite automaton with less than $|\text{rep}_U(a_X - 1)| - \iota_U(p_X)$ states accepting $\text{rep}_U(X)$. There exist words w, w_4 such that the greedy U-representation of $a_X - 1$ can be factorized as

$$\text{rep}_U(a_X - 1) = w w_4$$

with $|w| = |\text{rep}_U(a_X - 1)| - \iota_U(p_X)$. By the pumping lemma, w can be written $w_1 w_2 w_3$ with $w_2 \neq \varepsilon$ and for all $i \geq 0$,

$$w_1 w_2^i w_3 w_4 \in \text{rep}_U(X) \Leftrightarrow w_1 w_2 w_3 w_4 \in \text{rep}_U(X).$$

By minimality of a_X and p_X, either $a_X - 1 \in X$ and for all $n \geq 1$, $a_X + n p_X - 1 \notin X$, or $a_X - 1 \notin X$ and for all $n \geq 1$, $a_X + n p_X - 1 \in X$. Using the ultimate periodicity of $(U_i \mod p_X)_{i \geq 0}$, we observe that, because $|w_4| = \iota_U(p_X)$, for all $i \geq 0$, we have

$$\text{val}_U(w_1 w_2^{i \pi_U(p_X)} w_2 w_3 w_4) \equiv \text{val}_U(w_1 w_2 w_3 w_4) + i \, \text{val}_U(w_2^{\pi_U(p_X)} 0^{|w_2 w_3 w_4|}) \mod p_X.$$

Therefore repeating a factor of length multiple of $\pi_U(p_X)$ exactly p_X times does not change the value mod p_X and we get

$$\text{val}_U(w_1 w_2^{p_X \pi_U(p_X)} w_2 w_3 w_4) \equiv \text{val}_U(w_1 w_2 w_3 w_4) \mod p_X,$$

leading to a contradiction. $\square$

For the sake of completeness, we restate some well-known properties of ultimately periodic sets (see for instance [37] for a prologue on Pascal's machine for b-ary systems).

Lemma 19. *Let a, b be nonnegative integers and $U = (U_i)_{i \geq 0}$ be a linear numeration system. The language*

$$\text{val}_U^{-1}(a\mathbb{N} + b) = \{w \in A_U^* \mid \text{val}_U(w) \in a\mathbb{N} + b\} \subset A_U^*$$

is regular. In particular, if $\mathbb{N}$ is U-recognizable then a DFA accepting $\text{rep}_U(a\mathbb{N} + b)$ can be obtained efficiently and any ultimately periodic set is U-recognizable.

Before giving the proof, notice that for any integer $n \geq 0$, $\text{val}_U^{-1}(n) \setminus 0^+ A_U^*$ is a finite set of words $\{x_1, \ldots, x_{t_n}\}$ over A_U such that $\text{val}_U(x_i) = n$ for all $i = 1, \ldots, t_n$. This set contains in particular $\text{rep}_U(n)$.

Proof. Since regular sets are stable under finite modification (i.e., adding or removing a finite number of words in the language), we can assume that $0 \leq b < a$. The sequence $(U_i \mod a)_{i \geq 0}$ is ultimately periodic with preperiod $\ell = \iota_U(a)$ and period $p = \pi_U(a)$. It is an easy exercise to build a deterministic finite automaton $\mathcal{A}$ accepting the reversal of the words in $\{w \in A_U^* \mid \mathrm{val}_U(w) \in a\mathbb{N} + b\}$. The alphabet of the automaton is A_U. States are pairs (r, s) where $0 \leq r < a$ and $0 \leq s < \ell + p$. The initial state is $(0, 0)$. Final states are the ones with the first component equal to b. Transitions are defined as follows

$$\forall s < \ell + p - 1 : \ (r, s) \xrightarrow{j} (jU_s + r \mod a, \ s + 1)$$

$$(r, \ell + p - 1) \xrightarrow{j} (jU_{\ell + p - 1} + r \mod a, \ \ell),$$

for all $j \in A_U$. Notice that $\mathcal{A}$ does not check the greediness of the accepted words, the construction only relies on the U-numerical value of the words modulo a.

For the particular case, one has to consider the intersection of two regular languages $\mathrm{rep}_U(\mathbb{N}) \cap \mathrm{val}_U^{-1}(a\mathbb{N} + b)$. $\qquad\square$

Remark 20. In the previous statement, the assumption about the U-recognizability of $\mathbb{N}$ is of particular interest. Indeed, it is well-known that for an arbitrary linear numeration system, $\mathbb{N}$ is in general *not* U-recognizable. If $\mathbb{N}$ is U-recognizable, then U satisfies a linear recurrence relation [39], but the converse does not hold. Sufficient conditions on the recurrence relation that U satisfies for $\mathbb{N}$ to be U-recognizable are given in [25].

Our decision procedure will also make use of the following result.

Lemma 21. *Let $U = (U_i)_{i \geq 0}$ be an increasing sequence satisfying a linear recurrence relation of order k of the kind (1). The following assertions are equivalent:*

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

In particular, if $a_k = \pm 1$, then $\lim_{m \to +\infty} N_U(m) = +\infty$.

Proof. It is enough to show that (ii) implies (i). Let the prime decomposition of $|a_k|$ be $|a_k| = p_1^{u_1} \cdots p_r^{u_r}$ with $u_1, \ldots, u_r > 0$. It is obvious that if $m = p_1^{v_1} \cdots p_r^{v_r} c$ with $v_1, \ldots, v_r \geq 0$ and $\gcd(a_k, c) = 1$ then

$$\pi_U(m) = \mathrm{lcm}\{\pi_U(p_1^{v_1}), \ldots, \pi_U(p_r^{v_r}), \pi_U(c)\}.$$

Notice that m tends to infinity if and only if at least one of the v_j's or c tends to infinity.

Assume first that for some $j \in \{1, \ldots, r\}$, $v_j \to +\infty$. By assumption we have that $\lim_{v_j \to +\infty} N_U(p_j^{v_j}) = +\infty$ and by Remark 17, we get that $\lim_{v_j \to +\infty} \pi_U(p_j^{v_j}) = +\infty$. Therefore $\pi_U(m)$ takes values larger than any constant by considering an integer m which is divisible by a sufficiently large power of p_j. Again using Remark 17, the same conclusion holds for $N_U(m)$.

Let $C = \{c_0 < c_1 < c_2 < \cdots\}$ be the set of natural numbers prime to a_k. For all $c \in C$, the sequence $(U_i \mod c)_{i \geq 0}$ is ultimately periodic but it is even purely periodic. Indeed, for all $i \geq 0$, U_{i+k} is determined by the k previous terms $U_{i+k-1}, \ldots, U_i$. But since $\gcd(a_k, c) = 1$, a_k is invertible modulo c and for all $i \geq 0$, $U_i \mod c$ is also determined by the k following terms $U_{i+1}, \ldots, U_{i+k}$. By definition of $N_U(c)$, the sequence $(U_i \mod c)_{i \geq 0}$ takes exactly $N_U(c)$ different values because any term appears infinitely often. Let α be the function mapping $m \in \mathbb{N}$ onto the smallest index $\alpha(m)$ such that $U_{\alpha(m)} \geq m$. Since U is increasing, α is non-decreasing and $\lim_{m \to +\infty} \alpha(m) = +\infty$. From this last observation and from the

pure periodicity of $(U_i \mod c)_{i \geq 0}$, it follows that for all $c \in C$, $N_U(c) \geq \alpha(c)$ because $U_0 < \cdots < U_{\alpha(c)-1} < c$. Consequently, we get

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

Any large enough integer m contains either a large power of some p_j or some large c prime to a_k and consequently (i) holds. $\qquad \square$

Theorem 22. *Let $U = (U_i)_{i \geq 0}$ be a linear numeration system such that $\mathbb{N}$ is U-recognizable, satisfying condition* (2)*. Assume that*

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

Then it is decidable whether or not a U-recognizable set is ultimately periodic.

Proof. The sequence U satisfies a recurrence relation of order k of the kind (1). Let the prime decomposition of $|a_k|$ be $|a_k| = p_1^{u_1} \cdots p_r^{u_r}$ with $u_1, \ldots, u_r > 0$. Consider a DFA $\mathcal{A}$ with d states accepting a U-recognizable set $X \subseteq \mathbb{N}$. Assume that X is ultimately periodic with a period

$$p_X = p_1^{v_1} \cdots p_r^{v_r} c$$

where $\gcd(a_k, c) = 1$ and $v_1, \ldots, v_r \geq 0$.

Since $\gcd(a_k, c) = 1$, with the same reasoning as in the proof of the previous lemma, the sequence $(U_i \mod c)_{i \geq 0}$ is purely periodic. Therefore, $U_0 = 1$ appears infinitely often in $(U_i \mod c)_{i \geq 0}$. Since c is a divisor of p_X, we can use Proposition 15 and get $c \leq d$.

By Proposition 13, we get $N_U(p_X) \leq d$. Let $j \in \{1, \ldots, r\}$. Using Remark 17, we obtain

$$N_U(p_j^{v_j}) \leq \pi_U(p_j^{v_j}) \leq \pi_U(p_X) \leq (N_U(p_X))^k \leq d^k.$$

The assumption $\lim_{m \to +\infty} N_U(m) = +\infty$ implies that $\lim_{v \to +\infty} N_U(p_j^v) = +\infty$. Observe that $N_U(p_j^v) \leq N_U(p_j^w)$ whenever $v \leq w$. Consequently the exponent v_j occurring in the decomposition of p_X is bounded by s_j where s_j is the smallest integer such that for all $v \geq s_j$, $N_U(p_j^v) > d^k$. This bound s_j can be effectively computed as follows. For any v, $N_U(p_j^v)$ can be found in a finite number of operations by inspecting the first values of $(U_i \mod p_j^v)_{i \geq 0}$ and looking for two identical k-tuples made of k consecutive elements. Once the period is determined, one immediately gets the values that are repeated infinitely often. Since the map $v \mapsto N_U(p_j^v)$ is non-decreasing, one has to compute $N_U(p_j) \leq N_U(p_j^2) \leq \cdots$ until finding the first value s_j such that $N_U(p_j^{s_j}) > d^k$.

If X is ultimately periodic, then the admissible periods are bounded by the constant

$$P = p_1^{s_1} \cdots p_r^{s_r} d$$

which is effectively computable. Then, using Proposition 18, the admissible preperiods a_X must satisfy

$$|\operatorname{rep}_U(a_X - 1)| \leq d + \max_{p \leq P}(\iota_U(p))$$

where $|\operatorname{rep}_U(a)| \leq |\operatorname{rep}_U(b)|$ whenever $a \leq b$. This last observation shows that a bound on the admissible preperiods of X can be given effectively.

Consequently the sets of admissible preperiods and periods that we have to check are finite. For each pair (a, p) of admissible preperiods and periods, there are at most $2^a 2^p$ distinct ultimately periodic sets. Thanks to Lemma 19, one can build an automaton for each of them and then compare the language L accepted by this automaton with $\operatorname{rep}_U(X)$. Recall that testing whether $L \setminus \operatorname{rep}_U(X) = \emptyset$ and $\operatorname{rep}_U(X) \setminus L = \emptyset$ is decidable algorithmically. $\qquad \square$

In view of the previous result, it is natural to characterize linear recurrence sequences U such that $\lim_{m \to +\infty} N_U(m) = +\infty$. It is exactly the aim of Section 3. For instance, we have the immediate special case that was treated in [8].

Corollary 23. *Let $U = (U_i)_{i \geq 0}$ be a linear numeration system such that $\mathbb{N}$ is U-recognizable and satisfying a recurrence relation of order k of the kind* (1) *with $a_k = \pm 1$ and condition* (2). *It is decidable whether or not a U-recognizable set is ultimately periodic.*

Proof. It is enough to observe that if $a_k = \pm 1$, then $\lim_{m \to +\infty} N_U(m) = +\infty$, see Lemma 21. □

Example 24. Let $\alpha \in (0,1)$ be irrational, with simple continued fraction $\alpha = [1, a_1, a, a_3, a, \ldots]$, that is, $a_{2i} = a$ $(i \geq 1)$, where $a, a_{2i-1} \in \mathbb{N} \setminus \{0\}$ for all $i \geq 1$. The numerators of its even-indexed convergents are given by the recurrence $U_{2i} = (aa_{2i-1} + 2)U_{2i-2} - U_{2i-4}$ $(i \geq 1)$ with initial conditions $U_{-2} = 1 - a$, $U_0 = 1$. It was shown in [16] that every positive integer n has a unique representation of the form $n = \sum_{i \geq 0} d_{2i}U_{2i}$, where the digits satisfy $0 \leq d_{2i} \leq aa_{2i+1} + 1$, and the additional condition: If for some $0 \leq k < \ell$ the digits d_{2k} and $d_{2\ell}$ attain their maximum values, then there exists j satisfying $k < j < \ell$ such that $d_{2j} < aa_{2j+1}$. It was recently employed in [38], where it was dubbed ℓ-sequences. The special case $a = a_{2i-1} = 1$ for all $i \geq 1$ is an exotic ternary numeration system since then $0 \leq d_{2i} \leq 2$. The additional condition then states that between any two digits 2 there must be a digit 0. Also $\alpha = (1 + \sqrt{5})/2$ is the golden ratio and the U_{2i} are the even-indexed Fibonacci numbers. This special case was used to investigate irregularities of distribution of sequences [10], [11]. In [19] it was used to compress sparse bit-strings, in [17] for answering a combinatorial question raised in [6], and in [18] for providing a polynomial-time algorithm for combinatorial games. For all these numeration systems Problem 1 is decidable, since the coefficient of U_{2i-4} is -1. The same conclusion holds for the numeration system based on the denominators of the even convergents, also given in [16].

Remark 25. We have thus obtained a decision procedure for our Problem 1 when $\lim_{m \to +\infty} N_U(m) = +\infty$ and in particular when the coefficient a_k occurring in (1) is equal to ± 1. On the other hand, whenever $\gcd(a_1, \ldots, a_k) = g \geq 2$, for all $n \geq 1$ and for all i large enough, we have $U_i \equiv 0 \mod g^n$ and the assumption about $N_U(m)$ in Theorem 22 does not hold. Indeed, the only value taken infinitely often by the sequence $(U_i \mod g^n)_{i \geq 0}$ is 0, so $N_U(m)$ equals 1 for infinitely many values of m. Notice in particular, that the same observation can be made for the usual b-ary system $(b \geq 2)$ where the only value taken infinitely often by the sequence $(b^i \mod b^n)_{i \geq 0}$ is 0, for all $n \geq 1$.

To conclude this section, we make a small digression. We show how to use a result of Engstrom about preperiods [14] to get some special linear numeration systems where $\lim_{m \to +\infty} N_U(m) = +\infty$. In Engstrom's paper the problem of finding a general period for a given recurrence relation modulo m for *any* initial conditions is considered. Notice that in [40], M. Ward considers the problem where the initial conditions are fixed and then the period modulo m has to be determined.

Theorem 26. [14, Theorem 9] *Let $U = (U_i)_{i \geq 0}$ be a linear recurrence sequence of order k of the kind* (1) *and p be a prime divisor of a_k. If there exists $s(p) < k$ such that $a_k, \ldots, a_{k-s(p)+1} \equiv 0 \pmod{p}$ and $a_{k-s(p)} \not\equiv 0 \pmod{p}$, then $\iota_U(p^v) \leq vs(p)$.*

Remark 27. Assume that we are dealing with a linear numeration system $U = (U_i)_{i \geq 0}$ satisfying (1) and that the assumptions of the previous theorem hold for all prime divisors p of a_k (which is equivalent to the fact that $\gcd(a_1, \ldots, a_k) = 1$).

Let $\chi_U(x)$ be the characteristic polynomial of U as defined in (4). Assume that $\beta > 1$ is a root of multiplicity $\ell \geq 1$ of $\chi_U(x)$ satisfying:

- for any other root $\gamma \in \mathbb{C}$ of $\chi_U(x)$, $|\gamma| < \beta$,
- $\beta < p^{1/s(p)}$ for all prime divisors p of a_k.

There exists some constant c such that $U_i \sim c\,i^{\ell-1}\beta^i$. Let p be a prime divisor of a_k and $j_p(v)$ be the largest index j such that $U_j < p^v$. Let $t > s(p)$ be a real number such that $\beta < p^{1/t} < p^{1/s(p)}$. For v large enough, we have $U_{\lfloor vt \rfloor} < p^v$. Consequently, for v large enough, $j_p(v) \geq \lfloor vt \rfloor$. From the previous theorem, we have $\iota_U(p^v) \leq vs(p)$. Therefore for v large enough, $U_{\iota_U(p^v)} < \cdots < U_{j_p(v)}$ are the first terms of the periodic part of $(U_i \mod p^v)_{i \geq 0}$ and $N_U(p^v) \geq \lfloor vt \rfloor - vs(p) + 1$. This means that for all prime divisors p of a_k, $N_U(p^v) \to +\infty$ as $v \to +\infty$. Therefore, by Lemma 21, $\lim_{m \to +\infty} N_U(m) = +\infty$ and we can apply our decision procedure given by Theorem 22 (whenever $\mathbb{N}$ is U-recognizable).

Example 28. Consider the linear recurrence sequence given by $U_{i+3} = U_{i+1} + 3U_i$ for $i \geq 0$ and $U_i = i + 1$ for $i = 0, 1, 2$. The first terms of the sequence are

$$1, 2, 3, 5, 9, 14, 24, 41, 66, 113, 189, 311, 528, 878, 1461, 2462, 4095, \ldots.$$

With the above notation, $\beta \simeq 1.6717 < 3$ and the other two complex roots have modulus close to 1.34. We also have $s(3) = 1$. Thanks to Theorem 26, the preperiod $\iota_U(3^v)$ is bounded by v. On the other hand, we have $U_i \sim c\,\beta^i$ for some $c > 0$. Notice that $\beta < 3^{1/2} < 3$. Therefore for v large enough, $U_{2v} \sim c\,\beta^{2v} < 3^v$. Consequently, the elements $U_v < \cdots < U_{2v}$ appear in the periodic part. In the following table, these elements have been underlined.

v	preperiod	period
3	1, 2, 3	$(\underline{5, 9, 14, 24}, 14, 12, 5, 0, 14, 15, 14, 3, 5, 18, 14, 6, 14, 21)$
4	1, 2, 3, 5	$(\underline{9, 14, 24, 41}, 66, 32, 27, 68, 42, 68, 3, 32, 45, 41, 60, 14, \ldots)$
5	1, 2, 3, 5, 9	$(\underline{14, 24, 41, 66, 113, 189}, 68, 42, 149, 3, 32, 207, 41, 60, 176, \ldots)$

3. Linear Recurrence Sequences and Residue Classes

As was observed in Remark 25, since our approach to solve the decision problem is requiring that $\lim_{m \to +\infty} N_U(m) = +\infty$, it can only be applied to linear recurrence sequences (1) for which $\gcd(a_1, \ldots, a_k) = 1$. In this section, our aim is to determine which linear recurrence sequences U are such that $\lim_{m \to +\infty} N_U(m) = +\infty$. To that end, it is clear (see Lemma 21) that we have only to focus on the behavior of $N_U(p^v)$ for any prime p dividing a_k.

Throughout this section we let $(U_i)_{i \geq 0}$ be a linear recurrence sequence satisfying (1). We assume that $(U_i)_{i \geq 0}$ satisfies no recurrence of smaller order than k. It is well-known (this result is sometimes referred as Kronecker's theorem, see [28]) that under the assumption $a_k \neq 0$, this is equivalent to assume that

$$\det \begin{pmatrix} U_0 & \cdots & U_{k-1} \\ \vdots & & \vdots \\ U_{k-1} & \cdots & U_{2k-2} \end{pmatrix} \neq 0.$$

Let p be a prime number. We recap some background on the p-adic numbers (see for instance [23]). We can put an absolute value $|\cdot|_p$ on $\mathbb{Z}$ as follows. For each integer $n \neq 0$, we can write $n = p^v \ell$ with ℓ such that $\gcd(p, \ell) = 1$. We define $|n|_p = p^{-v}$ and $|0|_p = 0$. We note that this absolute value extends on $\mathbb{Q}$ by declaring $|a/b|_p = |a|_p/|b|_p$ for $a, b \in \mathbb{Z}$, $b \neq 0$. In particular, for all $a, b \in \mathbb{Q}$, $|a.b|_p = |a|_p.|b|_p$. Note that this absolute value is *non-Archimedian*; that is, it satisfies for all $a, b \in \mathbb{Q}$

$$|a + b|_p \leq \max\{|a|_p, |b|_p\}.$$

If we complete $\mathbb{Q}$ with respect to this absolute value, we obtain the field of *p-adic rationals*, which we denote by $\mathbb{Q}_p$. We can take the algebraic closure of $\mathbb{Q}_p$; the absolute value $|\cdot|_p$ extends to this algebraic closure. The algebraic closure is not complete, however. Completing this algebraic closure, we obtain a complete algebraically closed field $\mathbb{C}_p$ with absolute value $|\cdot|_p$, which restricts to the p-adic absolute value on $\mathbb{Q}$. The closed unit ball

$$\{x \in \mathbb{Q}_p : |x|_p \leq 1\}$$

is called the set of *p-adic integers* and we denote it by $\mathbb{Z}_p$. The ordinary integers are dense in $\mathbb{Z}_p$. The p-adic rationals can be viewed as the formal expressions of the form

$$c_{-N}p^{-N} + \cdots + c_{-1}p^{-1} + c_0 + c_1 p + c_2 p^2 + \cdots$$

where $c_j \in \{0, \ldots, p-1\}$, $N \in \mathbb{Z}$. The p-adic integers are identified with formal expressions involving only non-negative powers of p.

We let

$$(4) \qquad \chi_U(x) = x^k - a_1 x^{k-1} - \cdots - a_k$$

be the characteristic polynomial of $(U_i)_{i \geq 0}$ and

$$(5) \qquad P_U(x) = x^k \chi_U(1/x) = 1 - a_1 x - \cdots - a_k x^k.$$

For $a_k \neq 0$, observe that if $\alpha_1, \ldots, \alpha_s$ are the roots of χ_U, then the reciprocals $1/\alpha_1, \ldots, 1/\alpha_s$ are exactly the roots of P_U.

Remark 29. If $(U_i)_{i \in \mathbb{N}}$ satisfies (1) with initial conditions $U_0, \ldots, U_{k-1}$, then the ordinary power series generating function is a rational function:

$$\mathsf{U}(x) := \sum_{i \geq 0} U_i x^i = \frac{\sum_{i=0}^{k-1} U_i x^i - \sum_{i+j<k} a_i U_j x^{i+j}}{P_U(x)}.$$

The converse also holds. The sequence of Taylor coefficients of any rational function $R(x)/P_U(x)$ where $R(x)$ is a polynomial of degree less than k, satisfies (1). See for instance [15, p. 6].

Our goal is to prove the following result. Notice that similar developments can be found in [34].

Theorem 30. *We have $N_U(p^v) \not\to +\infty$ 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:*
  (i) $B(x) \equiv 1 \pmod{p\mathbb{Z}[x]}$;
  (ii) $A(x)$ *has no repeated roots and all its roots are roots of unity.*
In that case, we have moreover $A(0) = B(0) = 1$.

Proof. We note that one direction is fairly simple. Assume that $P_U(x)$ has such a factorization $P_U(x) = A(x)B(x)$. From (ii), there is a natural number d such that $A(x)|(x^d - 1)$. In view of Remark 29, there exist some polynomials $Q(x), R(x)$ such that

$$(x^d - 1)\sum_{i \geq 0} U_i x^i \ = \ \frac{(x^d - 1)Q(x)}{P_U(x)} \ = \ \frac{(x^d - 1)Q(x)}{A(x)B(x)} \ = \ \frac{R(x)}{B(x)}.$$

By assumption (i), there exists some integer polynomial $B_1(x)$ such that $B(x) = 1 - pB_1(x)$. Hence

$$(x^d - 1)\sum_{i \geq 0} U_i x^i \ = \ \frac{R(x)}{1 - pB_1(x)} \ = \ \sum_{i \geq 0} p^i R(x) B_1(x)^i.$$

In particular, for any fixed integer v, $(x^d-1)\sum_{i\geq 0} U_i x^i$ is congruent to a polynomial (mod p^v). This means that $U_{i+d} \equiv U_i \pmod{p^v}$ for all i sufficiently large. In particular, there are at most d values which can occur infinitely often mod p^v; that is, $N_U(p^v) \leq d$ for every v.

To do the other direction is a little more work and we use p-adic methods. We first note that $v \mapsto N_U(p^v)$ is a non-decreasing function, i.e.,

$$(6) \qquad\qquad N_U(p^w) \geq N_U(p^v) \text{ whenever } w \geq v.$$

In particular, if $N(p^v) \not\to +\infty$ then there is some d such that $N_U(p^v) = d$ for all v sufficiently large. We can pick integers $a_{1,v}, \ldots, a_{d,v}$ such that if $U_i \equiv a \pmod{p^v}$ for infinitely many i, then $a \equiv a_{j,v} \pmod{p^v}$ for some $j \in \{1, \ldots, d\}$. Since

$$\{(a_{1,w} \bmod p^v), \ldots, (a_{d,w} \bmod p^v)\} = \{(a_{1,v} \bmod p^v), \ldots, (a_{d,v} \bmod p^v)\}$$

for $w \geq v$, there is no loss of generality to assume that

$$a_{j,w} \equiv a_{j,v} \pmod{p^v}$$

for $w \geq v$, $1 \leq j \leq d$. It follows that for $1 \leq j \leq d$ the sequence $(a_{j,v})_{v \geq 1}$ is Cauchy in $\mathbb{Z}_p$. Thus there exist $b_1, \ldots, b_d \in \mathbb{Z}_p$ such that for $1 \leq j \leq d$,

$$(7) \qquad\qquad a_{j,v} \to b_j \text{ as } v \to +\infty.$$

Let

$$V_i \;=\; \prod_{j=1}^{d} (U_i - b_j) \in \mathbb{Z}_p.$$

Note that since the set of linear recurrence sequences is closed under (Hadamard) product, sum, and scalar multiplication (see for instance [15]), the sequence $(V_i)_{i\geq 0}$ satisfies a linear recurrence over $\mathbb{Z}_p$. By construction, V_i is eventually in $p^v \mathbb{Z}_p$ for any fixed v, since any values of U_i that are not congruent to one of $b_1, \ldots, b_d$ mod p^v can only occur finitely many times. That is for any v, $|V_i|_p \leq p^{-v}$ for i sufficiently large. Hence $|V_i|_p \to 0$ as $i \to +\infty$. Since $(V_i)_{i\geq 0}$ satisfies a linear recurrence, the power series

$$\mathsf{V}(x) := \sum_{i\geq 0} V_i x^i$$

is a rational power series in $\mathbb{Q}_p(x)$. Moreover, $\mathsf{V}(x)$ converges on the closed unit disc $\mathbb{Z}_p$, since $|V_i|_p \to 0$ (in non-Archimedian fields this is enough to guarantee convergence: a series $\sum_{i\geq 0} \gamma_i$ converges in $\mathbb{Q}_p$ if and only if $\lim_{i\to +\infty} |\gamma_i|_p = 0$). Since $\mathsf{V}(x)$ is a rational series and it converges on the unit disc, its poles $\beta_1, \ldots, \beta_r \in \mathbb{C}_p$ must satisfy $|\beta_j|_p > 1$, for $1 \leq j \leq r$.

To finish the proof, we will make use of Lemma 31 stated below. For the sake of clarity, we have separated this technical result from this proof. We note that the statement of this lemma is very close to what we already have, but it makes the additional assumption that the poles of $\mathsf{U}(x)$ generate a free Abelian subgroup. In general, the poles of $\mathsf{U}(x)$ generate a finitely generated Abelian subgroup of $\mathbb{C}_p^\times$. From the so-called *fundamental theorem of finitely generated Abelian groups* (see for instance [5, p. 141]), this group is isomorphic to $\mathbb{Z}^e \times T$, for some finite Abelian group T and integer $e \geq 0$.

Let us show how to get rid of the torsion group T to be able to invoke Lemma 31. Let $a = \#T$. For $0 \leq b < a$, instead of taking the sequence $(U_i)_{i\geq 0}$, consider the sequence $(U_i^{(b)})_{i\geq 0} := (U_{ai+b})_{i\geq 0}$. This latter sequence satisfies a linear recurrence

and the poles of the generating function $\mathsf{U}^{(b)}(x)$ of $(U_i^{(b)})_{i\geq 0}$ are the a'th powers[1] of the poles of $\mathsf{U}(x)$. Consequently, the poles of $\mathsf{U}^{(b)}(x)$ generate a finitely generated torsion-free Abelian group, which is necessarily a free Abelian group.

Since the poles $\beta_1, \ldots, \beta_r \in \mathbb{C}_p$ of $\mathsf{V}(x)$ satisfy $|\beta_j|_p > 1$, for $1 \leq j \leq r$, with the same reasoning, we get that the poles of the rational function

$$\mathsf{V}^{(b)}(x) = \sum_{i\geq 0} V_{ai+b}\, x^i$$

are a'th powers of $\beta_1, \ldots, \beta_r$ and for $1 \leq j \leq r$, $|\beta_j^a|_p = |\beta_j|_p^a > 1$. We can invoke Lemma 31 applied to the sequence $(U_i^{(b)})_{i\geq 0}$ (it satisfies some linear recurrence sequence with integer coefficients), and we deduce that any pole γ of $\mathsf{U}^{(b)}(x)$ satisfies either $|\gamma|_p > 1$ or $\gamma = 1$. In particular, the distinct poles $1/\alpha_1, \ldots, 1/\alpha_s$ of $\mathsf{U}(x)$, with $\alpha_1, \ldots, \alpha_s \in \mathbb{C}_p$, being a'th roots of the poles of $\mathsf{U}^{(b)}(x)$, are either such that $|1/\alpha_j|_p > 1$ or roots of unity. By minimality of the order k of the recurrence satisfied by $(U_i)_{i\geq 0}$, the poles of $\mathsf{U}(x)$ are precisely the roots of $P_U(x)$. We factor

$$P_U(x) = (1 - \delta_1 x) \cdots (1 - \delta_k x)$$

(each δ_j is one of $\alpha_1, \ldots, \alpha_s$, although they may be repeated). Let us factor $P_U(x)$ as $A(x)B(x)$ where

$$A(x) = \prod_{\{j:|\delta_j|_p=1\}} (1 - \delta_j x) \quad \text{and} \quad B(x) = \prod_{\{j:|\delta_j|_p<1\}} (1 - \delta_j x).$$

By assumption $P_U(x) \in \mathbb{Z}[x]$. Moreover if K is a splitting field of $P_U(x)$ over $\mathbb{Q}$ then any automorphism of K must permute the set of δ_j with $|\delta_j|_p < 1$, since the automorphism permutes the entire set of δ_j's and it must send roots of unity to roots of unity. Thus $B(x)$ is a rational polynomial, since it is fixed by every automorphism of K. Note that if $n > 0$, then the coefficient of x^n in $B(x)$ is given by a sum of products of n elements in $\{\delta_j : |\delta_j|_p < 1\}$. The set of algebraic integers is a subring of $\mathbb{C}_p$ and the only rationals that are algebraic integers are in fact integers. Since the δ_j's are algebraic integers, $B(x)$ is thus an integer polynomial. Moreover, since the p-adic absolute value is non-Archimedian, the coefficient of x^n in $B(x)$, $n > 0$, has p-adic absolute value strictly less than 1. Note that an integer m satisfying $|m|_p < 1$ is necessarily a multiple of p. Hence $B(x) \equiv 1 \bmod p\mathbb{Z}[x]$.

Let us turn to the polynomial $A(x)$. The roots of $A(x)$ are roots of unity. Moreover, $A(x) \in \mathbb{Z}[x]$ by the same reasoning as before.

To finish the proof, we have one last thing to show: we need to know that $A(x)$ has no repeated roots. To do this, we show that the poles of $\mathsf{U}(x)$ that are roots of unity are simple. Recall that $1/\alpha_1, \ldots, 1/\alpha_s$ are the distinct poles of $\mathsf{U}(x)$. We may assume that there exists $t \geq 0$ such that $\alpha_1, \ldots, \alpha_t$ are roots of unity and that for $j = t+1, \ldots, s$, $|\alpha_j|_p < 1$. Then there exist polynomials $q_j \in \mathbb{C}_p[x]$, $j = 1, \ldots, s$, such that we have for all i,

$$U_i = \sum_{j=1}^{s} q_j(i)\,\alpha_j^i = \underbrace{\sum_{j=1}^{t} q_j(i)\,\alpha_j^i}_{:=T_i} + \underbrace{\sum_{j=t+1}^{s} q_j(i)\,\alpha_j^i}_{:=W_i}.$$

Since for $j = t+1, \ldots, s$, $|\alpha_j|_p < 1$, we get that $|U_i - T_i|_p = |W_i|_p \to 0$ as $i \to +\infty$. Indeed, for any j, $\{|q_j(i)|_p : i \in \mathbb{N}\}$ is bounded by a constant. Since for $1 \leq j \leq t$,

---

[1]If we consider the exponential sum $U_i = \sum_{j=1}^{s} q_j(i)\,\alpha_j^i$ where the q_j's are polynomials in $\mathbb{C}_p[x]$ and the α_j's are the roots of the characteristic polynomial of $(U_i)_{i\geq 0}$, i.e., the reciprocals of the poles of $\mathsf{U}(x)$, then we obtain that $U_i^{(b)} = U_{ai+b} = \sum_{j=1}^{s} \alpha_j^b P_j(ai+b)\,(\alpha_j^a)^i$. Hence the poles of $\mathsf{U}^{(b)}(x)$ are the a'th powers of the poles of $\mathsf{U}(x)$.

α_j is a root of unity, there exists a natural number a such that $\alpha_j^a = 1$ for all $j \in \{1, \ldots, t\}$. As before, we let $T_i^{(b)} = T_{ai+b}$ for $0 \le b < a$. Thus

$$T_i^{(b)} = \sum_{j=1}^t q_j(ai+b)\,\alpha_j^{ai+b} = \sum_{j=1}^t \alpha_j^b\, q_j(ai+b)$$

is a polynomial with coefficients in $\mathbb{C}_p$ denoted by $g_b(i)$.

Let $\epsilon > 0$. By definition of the b_j's (see (7) above) for all large enough i, there exists $\ell(i) \in \{1, \ldots, d\}$ such that $|U_i - b_{\ell(i)}|_p < \epsilon$. Since $|U_i - T_i|_p \to 0$ as $i \to +\infty$, we get

$$|(T_i - b_1) \cdots (T_i - b_d)|_p \le \prod_{j=1}^d (|T_i - U_i|_p + |U_i - b_{\ell(i)}|_p + |b_{\ell(i)} - b_j|_p) \to 0$$

as $i \to +\infty$ because every factor is bounded by a constant and one tends to zero. Thus for $0 \le b < a$, as $T_i^{(b)} = g_b(i)$, we have

$$|(g_b(i) - b_1) \cdots (g_b(i) - b_d)|_p \to 0 \text{ as } i \to +\infty.$$

Consider the polynomial defined by

$$h_b(i) := (g_b(i) - b_1) \cdots (g_b(i) - b_d).$$

Let n_0 be a natural number. Notice that by the Binomial Theorem, we get $|h_b(n_0 + p^v) - h_b(n_0)|_p \le Cp^{-v} \to 0$ as $v \to +\infty$, where C is the maximum of the p-adic absolute values of the coefficients of h_b. So since $|h_b(i)|_p \to 0$ as $i \to +\infty$, we get[2] that $h_b(n_0) = 0$. Thus for all integer i each $h_b(i) = 0$, $0 \le b < a$, and so each g_b is a constant polynomial (this is a consequence of the *Fundamental Theorem of Algebra*, $h_b(i) = 0$ for all i implies that g_b takes infinitely often the same value amongst $b_1, \ldots, b_d$ so it is constant). It follows that T_i is periodic:

$$\forall i \ge 0,\ T_{i+a} = T_i.$$

So $(x^a - 1)\mathsf{U}(x)$ has no poles on the closed unit disc so in particular on the unit circle, as

$$|U_{i+a} - U_i|_p \le |T_{i+a} - T_i|_p + |W_{i+a} - W_i|_p \to 0$$

as $i \to +\infty$. We know that $\mathsf{U}(x) = Q(x)/P_U(x)$ with Q and P_U relatively prime (by minimality assumption on order of recurrence satisfied by $(U_i)_{i \ge 0}$). Consequently, $A(x)$ divides $(x^a - 1)$ because $(x^a - 1)Q(x)/P_U(x)$ has no poles on the unit circle. This shows that it has no repeated roots, completing the proof. $\square$

Let us consider the technical lemma used in the previous proof.

Lemma 31. *Let $(U_i)_{i \ge 0}$ be an integer linear recurrence sequence satisfying (1) and let $\mathsf{U}(x) = \sum_{i \ge 0} U_i x^i \in \mathbb{Z}_p[[x]]$ be the corresponding rational power series. Suppose that the multiplicative subgroup of $\mathbb{C}_p^\times$ generated by the (finitely many) poles of $\mathsf{U}(x)$ is a free Abelian group and let $V_i = \prod_{j=1}^d (U_i - b_j)$ for $i \ge 0$ where $b_1, \ldots, b_d \in \mathbb{Z}_p$. If the rational power series $\mathsf{V}(x) = \sum_{i \ge 0} V_i x^i \in \mathbb{Z}_p[[x]]$ has poles $\beta_1, \ldots, \beta_r \in \mathbb{C}_p$ satisfying $|\beta_j|_p > 1$ for $1 \le j \le r$, then every pole $\gamma \in \mathbb{C}_p$ of $\mathsf{U}(x)$ either satisfies $|\gamma|_p > 1$ or $\gamma = 1$.*

Proof. Let $1/\alpha_1, \ldots, 1/\alpha_s$ be the distinct poles of $\mathsf{U}(x)$ with $\alpha_1, \ldots, \alpha_s \in \mathbb{C}_p$. Note that 0 cannot be a pole of U. We first claim that $|\alpha_j|_p \le 1$ for $1 \le j \le s$. To see this, notice that to be a pole, each $1/\alpha_j$ must satisfy $P_U(1/\alpha_j) = 0$. This means that

$$1 - a_1/\alpha_i - \cdots - a_k/\alpha_i^k = 0.$$

---

[2]Indeed, $|h_b(n_0)|_p \le |h_b(n_0) - h_b(n_0 + p^v)|_p + |h_b(n_0 + p^v)|_p$ and both terms tend to 0 as $v \to +\infty$.

Consequently, for all $j \in \{1, \ldots, s\}$, we have

$$\left| a_1/\alpha_j + \cdots + a_k/\alpha_j^k \right|_p = |1|_p = 1.$$

By the non-Archimedian property we have $|a_\ell/\alpha_j^\ell|_p \geq 1$ for some $\ell \in \{1, \ldots, k\}$. Thus $|\alpha_j|_p^\ell \leq |a_\ell|_p$. Since $a_\ell \in \mathbb{Z}$, $|a_\ell|_p \leq 1$. This gives the claim and so $|\alpha_j|_p \leq 1$ for $1 \leq j \leq s$. We may assume that $|\alpha_1|_p = \cdots = |\alpha_t|_p = 1$ and $|\alpha_j|_p < 1$ for $0 \leq t < j \leq s$. So, we know that the poles of U have p-adic absolute value ≥ 1.

Next there exist polynomials $q_1(x), \ldots, q_s(x) \in \mathbb{C}_p[x]$ such that

$$U_i \; = \; \sum_{j=1}^s q_j(i)\alpha_j^i.$$

Moreover, define

$$V_i := (U_i - b_1) \cdots (U_i - b_d) = c_d U_i^d + c_{d-1} U_i^{d-1} + \cdots + c_0$$

for some $c_0, \ldots, c_{d-1}, c_d = 1 \in \mathbb{Z}_p$. Hence by Multinomial Theorem, we have

$$\begin{aligned}
V_i \; &= \; \sum_{j=0}^d c_j \sum_{\substack{j_1 + \cdots + j_s = j \\ j_1, \ldots, j_s \geq 0}} \binom{j}{j_1, \ldots, j_s} \prod_{\ell=1}^s (q_\ell(i)\alpha_\ell^i)^{j_\ell} \\
&= \; \sum_{j=0}^d c_j \sum_{\substack{j_1 + \cdots + j_s = j \\ j_1, \ldots, j_s \geq 0}} \binom{j}{j_1, \ldots, j_s} \prod_{\ell=1}^s q_\ell(i)^{j_\ell} \left(\prod_{\ell=1}^s \alpha_\ell^{j_\ell} \right)^i.
\end{aligned}$$

Since the roots of the characteristic polynomial are the reciprocals of the poles of the corresponding rational power series, it follows that the set of poles of $\mathsf{V}(x)$ is contained in the set

$$\left\{ \prod_{\ell=1}^s \alpha_\ell^{-j_\ell} \; : \; j_1, \ldots, j_s \geq 0, \; j_1 + \cdots + j_s \leq d \right\}.$$

By assumption, the poles of $\mathsf{V}(x)$ all have p-adic absolute value strictly greater than 1. Note that $\left| \prod_{\ell=1}^r \alpha_\ell^{-j_\ell} \right|_p = \left| \prod_{\ell=1}^t \alpha_\ell^{-j_\ell} \right|_p \left| \prod_{\ell=t+1}^r \alpha_\ell^{-j_\ell} \right|_p$ is > 1 if and only if $j_\ell > 0$ for some $\ell > t$. Therefore, we can conclude that the possible poles of V supported only by products on $\alpha_1, \ldots, \alpha_t$ do not occur. So for instance, α_1^d is not a pole of V.

Let G denote the multiplicative subgroup of $\mathbb{C}_p^\times$ generated by $\alpha_1, \ldots, \alpha_t$. By assumption, G is a subgroup of a finitely generated free Abelian group, and hence $G \cong \mathbb{Z}^e$, for some natural number $e \geq 0$. To conclude the proof, it is enough to show that $e = 0$. Because in that case, one can conclude that the only possible pole γ of U such that $|\gamma|_p = 1$ is 1. To see this, suppose that $e > 0$ and let $\gamma_1, \ldots, \gamma_e$ be generators for a free Abelian group of rank e. Then for $1 \leq i \leq t$, we can write

$$\alpha_i = \prod_{j=1}^e \gamma_j^{b_{i,j}},$$

where $b_{i,j}$ are integers. We relabel if necessary so that

- $|b_{1,1}| = \max\{|b_{i,1}|\} > 0$;
- $|b_{1,j}| = \max\{|b_{i,j}| \; : \; b_{i,\ell} = b_{1,\ell} \text{ for } \ell < j\}$.

By construction, α_1^d cannot be written as a different word in $\alpha_1, \ldots, \alpha_t$ of length at most d. Then the expression for V_i above has an occurrence of

$$c_d q_1(i)^d \alpha_1^{di}$$

that cannot be canceled by any other pole, by our selection of α_1. Consequently, α_1^d should be a pole of V which contradicts the conclusion of the previous paragraph. So, $e = 0$ and the result follows.   □

As it was stressed in the introduction of this section, we are now able to check, whether or not the number of values taken infinitely often by a linear recurrence sequence modulo p^v tends to infinity as $v \to +\infty$. Let us consider the following example.

Example 32. Consider the linear recurrence sequence $(U_i)_{i\geq 0}$ given in [20] and defined by $U_{i+4} = 3U_{i+3} + 2U_{i+2} + 3U_i$ for $i \geq 0$ and $U_i = i+1$ for $i = 0, \ldots, 3$. As shown in [20], addition within this linear numeration system is not computable by a finite automaton. Nevertheless, we can show by applying the previous theorem that $N_U(3^v) \to +\infty$ as $v \to +\infty$. One has

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

and it is not difficult to see that this polynomial cannot be factorized as $A(x)B(x)$ with two factors satisfying hypotheses of Theorem 30. This example shows that our decision procedure given by Theorem 22 can take care of numeration systems not handled by [3, 30, 32].

Example 33. Consider the recurrence relation

$$U_{i+5} = 6U_{i+4} + 3U_{i+3} - U_{i+2} + 6U_{i+1} + 3U_i, \ \forall i \geq 0.$$

With the above notation, $\chi_U(x) = x^5 - 6x^4 - 3x^3 + x^2 - 6x - 3$ and

$$P_U(x) = 1 - 6x - 3x^2 + x^3 - 6x^4 - 3x^5 = \underbrace{(x^3 + 1)}_{A(x)} \underbrace{(-3x^2 - 6x + 1)}_{B(x)}.$$

With the initial conditions $U_i = i+1$ for $i = 0, \ldots, 4$, the corresponding sequence does not satisfy any relation of shorter length as

$$\det \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 54 \\ 3 & 4 & 5 & 54 & 359 \\ 4 & 5 & 54 & 359 & 2344 \\ 5 & 54 & 359 & 2344 & 15129 \end{pmatrix} = 8458240 \neq 0.$$

Even if the gcd of the coefficients of the recurrence is 1, since P_U satisfies the assumptions of Theorem 30 for $p = 3$, $N_U(3^v) \not\to +\infty$ as $v \to +\infty$. The following table gives the first values of $N_U(3^v)$.

v	period	$N_U(3^v)$
1	$(1, 0, 1, 2, 0, 2)$	3
2	$(4, 0, 1, 5, 0, 8)$	5
3	$(22, 9, 19, 5, 18, 8)$	6
4	$(49, 63, 19, 32, 18, 62)$	6
5	$(211, 225, 19, 32, 18, 224)$	6
⋮		⋮

4. A Decision Procedure for a Class of Abstract Numeration Systems

An *abstract numeration system* $S = (L, \Sigma, <)$ is given by an infinite regular language L over a totally ordered alphabet $(\Sigma, <)$ [29]. By enumerating the words of L in genealogical order, we get a one-to-one correspondence denoted rep_S between $\mathbb{N}$ and L. In particular, 0 is represented by the first word in L. The reciprocal map associating a word $w \in L$ to its index in the genealogically ordered language L is

denoted val_S. A set $X \subseteq \mathbb{N}$ of integers is *S-recognizable* if the language $\mathrm{rep}_S(X)$ over Σ is regular.

Let $S = (L, \Sigma, <)$ be an abstract numeration system built over an infinite regular language L having $\mathcal{M}_L = (Q_L, q_{0,L}, \Sigma, \delta_L, F_L)$ as minimal automaton. The transition function $\delta_L : Q_L \times \Sigma \to Q_L$ is extended on $Q_L \times \Sigma^*$ by $\delta_L(q, \varepsilon) = q$ and $\delta_L(q, aw) = \delta_L(\delta_L(q, a), w)$ for all $q \in Q_L$, $a \in \Sigma$ and $w \in \Sigma^*$. We denote by $\mathbf{u}_i(q)$ (resp. $\mathbf{v}_i(q)$) the number of words of length i (resp. $\leq i$) accepted from $q \in Q_L$ in $\mathcal{M}_L$. By classical arguments, the sequences $(\mathbf{u}_i(q))_{i \geq 0}$ (resp. $(\mathbf{v}_i(q))_{i \geq 0}$) satisfy the same homogenous linear recurrence relation for all $q \in Q_L$ (for details, see Remark 36).

In this section, we consider, with some extra hypothesis on the abstract numeration system, the following decidability question analogous to Problem 1.

Problem 2. Given an abstract numeration system S and a set $X \subseteq \mathbb{N}$ such that $\mathrm{rep}_S(X)$ is recognized by a (deterministic) finite automaton, is it decidable whether or not X is ultimately periodic, i.e., whether or not X is a finite union of arithmetic progressions ?

Abstract numeration systems are a generalization of positional numeration systems $U = (U_i)_{i \geq 0}$ for which $\mathbb{N}$ is U-recognizable.

Example 34. Take the language $L = \{\varepsilon\} \cup 1\{0, 01\}^*$ and assume $0 < 1$. Ordering the words of L in genealogical order: $\varepsilon, 1, 10, 100, 101, 1000, 1001, \ldots$ gives back the Fibonacci system.

The following example shows that the class of abstract numeration systems is strictly bigger than the class of linear numeration systems for which $\mathbb{N}$ is recognizable.

Example 35. Consider the language $L = \{\varepsilon\} \cup \{a, ab\}^* \cup \{c, cd\}^*$ and the ordering $a < b < c < d$ of the alphabet. If we order the first words in L we get

0	ε	5	cc	10	ccc	15	$aaba$	20	$ccdc$
1	a	6	cd	11	ccd	16	$abaa$	21	$cdcc$
2	c	7	aaa	12	cdc	17	$abab$	22	$cdcd$
3	aa	8	aab	13	$aaaa$	18	$cccc$	23	$aaaaa$
4	ab	9	aba	14	$aaab$	19	$cccd$	24	$aaaab$

Notice that there is no bijection $\nu : \{a, b, c, d\} \to \mathbb{N}$ between $\{a, b, c, d\}$ and a set of integers leading to a positional linear numeration system. Otherwise stated, a, b, c, d cannot be identified with usual "digits". Indeed, assume that there exists a sequence U of integers such that for all $x_1 \cdots x_n \in L$, with x_i in $\{a, b, c, d\}$ for all i, $\mathrm{val}_U(\nu(x_1) \cdots \nu(x_n)) = \mathrm{val}_S(x_1 \cdots x_n)$. Since $\mathrm{val}_S(a) = 1$ and $\mathrm{val}_S(c) = 2$ we get $U_0 = 1$, $\nu(a) = 1$ and $\nu(c) = 2$. Moreover, $\mathrm{val}_S(aa) = 3 = \nu(a)U_1 + \nu(a)U_0$, so $U_1 = 2$. Therefore, $\mathrm{val}_U(\nu(c)\nu(c)) = 2U_1 + 2U_0 = 6$ but $\mathrm{val}_S(cc) = 5$ getting a contradiction.
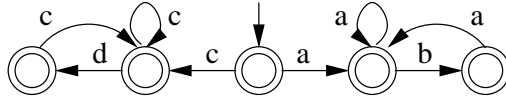

FIGURE 1. A DFA accepting L.

For all $i \geq 1$, we have $\mathbf{u}_i(q_{0,L}) = 2F_i$ and $\mathbf{u}_0(q_{0,L}) = 1$. Consequently, for $i \geq 1$,

$$\mathbf{v}_i(q_{0,L}) = 1 + \sum_{n=1}^{i} \mathbf{u}_n(q_{0,L}) = 1 + 2 \sum_{n=1}^{i} F_n.$$

Notice that for $i \geq 1$, $\mathbf{v}_i(q_{0,L}) - \mathbf{v}_{i-1}(q_{0,L}) = \mathbf{u}_i(q_{0,L}) = 2F_i$. Consequently, by definition of the Fibonacci sequence, we get for all $i \geq 3$,

$$\mathbf{v}_i(q_{0,L}) - \mathbf{v}_{i-1}(q_{0,L}) = (\mathbf{v}_{i-1}(q_{0,L}) - \mathbf{v}_{i-2}(q_{0,L})) + (\mathbf{v}_{i-2}(q_{0,L}) - \mathbf{v}_{i-3}(q_{0,L}))$$

and

$$\mathbf{v}_i(q_{0,L}) = 2\mathbf{v}_{i-1}(q_{0,L}) - \mathbf{v}_{i-3}(q_{0,L}), \text{ with } \mathbf{v}_0(q_{0,L}) = 1, \mathbf{v}_1(q_{0,L}) = 3, \mathbf{v}_2(q_{0,L}) = 7.$$

Remark 36. The computation given in the previous example to obtain a homogenous linear recurrence relation for the sequence $(\mathbf{v}_i(q_{0,L}))_{i \geq 0}$ can be carried on in general. Let $q \in Q_L$. The sequence $(\mathbf{u}_i(q))_{i \geq 0}$ satisfies a homogenous linear recurrence relation of order k whose characteristic polynomial is the characteristic polynomial of the adjacency matrix of $\mathcal{M}_L$. There exist $a_1, \ldots, a_k \in \mathbb{Z}$ such that for all $i \geq 0$,

$$\mathbf{u}_{i+k}(q) = a_1 \mathbf{u}_{i+k-1}(q) + \cdots + a_k \mathbf{u}_i(q).$$

Consequently, we have for all $i \geq 0$

$$\mathbf{v}_{i+k+1}(q) - \mathbf{v}_{i+k}(q) = \mathbf{u}_{i+k+1}(q) = a_1(\mathbf{v}_{i+k}(q) - \mathbf{v}_{i+k-1}(q)) + \cdots + a_k(\mathbf{v}_{i+1}(q) - \mathbf{v}_i(q)).$$

Therefore the sequence $(\mathbf{v}_i(q))_{i \geq 0}$ satisfies a homogenous linear recurrence relation of order $k+1$.

As shown by the following lemma, in an abstract numeration system, the different sequences $(\mathbf{u}_i(q))_{i \geq 0}$, for $q \in Q_L$, are replacing the single sequence $(U_i)_{i \geq 0}$ defining a positional numeration system as in Definition 1.

Lemma 37. [29] *Let $w = \sigma_1 \cdots \sigma_n \in L$. We have*

$$(8) \qquad \mathrm{val}_S(w) = \sum_{q \in Q_L} \sum_{i=1}^{|w|} \beta_{q,i}(w)\, \mathbf{u}_{|w|-i}(q)$$

where

$$(9) \qquad \beta_{q,i}(w) := \#\{\sigma < \sigma_i \mid \delta_L(q_{0,L}, \sigma_1 \cdots \sigma_{i-1}\sigma) = q\} + \mathbf{1}_{q,q_{0,L}}$$

for $i = 1, \ldots, |w|$.

Recall that $\mathbf{1}_{q,q'}$ is equal to 1 if $q = q'$ and it is equal to 0 otherwise.

Proposition 38. [29] *Let $S = (L, \Sigma, <)$ be an abstract numeration system built over an infinite regular language L over Σ. Any ultimately periodic set X is S-recognizable and a DFA accepting $\mathrm{rep}_S(X)$ can be effectively obtained.*

Recall that an automaton is *trim* if it is accessible and coaccessible (each state can be reached from the initial state and from each state, one can reach a final state).

Proposition 39. *Let $S = (L, \Sigma, <)$ be an abstract numeration system such that for all states q of the trim minimal automaton $\mathcal{M}_L = (Q_L, q_{0,L}, \Sigma, \delta_L, F_L)$ of L,*

$$\lim_{i \to +\infty} \mathbf{u}_i(q) = +\infty$$

and $\mathbf{u}_i(q_{0,L}) > 0$ for all $i \geq 0$. If $X \subseteq \mathbb{N}$ is an ultimately periodic set of period p_X, then any deterministic finite automaton accepting $\mathrm{rep}_S(X)$ has at least $\lceil N_{\mathbf{v}}(p_X)/\#Q_L \rceil$ states where $\mathbf{v} = (\mathbf{v}_i(q_{0,L}))_{i \geq 0}$.

Proof. Let a_X be the preperiod of X. Since for all states q of $\mathcal{M}_L$, we have $\lim_{i \to +\infty} \mathbf{u}_i(q) = +\infty$, there exists a minimal constant $J > 0$ such that $\mathbf{u}_J(q) \geq p_X$ for all $q \in Q_L$. Consider for any $i \geq 0$, the word

$$w_i = \mathrm{rep}_S(\mathbf{v}_i(q_{0,L})),$$

corresponding to the first word of length $i+1$ in the genealogically ordered language L. Consequently, for $i \geq J - 1$, w_i is factorized as $w_i = a_i b_i$ with $|b_i| = J$ and we define $q_i := \delta_L(q_{0,L}, a_i)$. Notice that b_i is the smallest word of length J accepted from q_i. By definition of J, from each q_i, there are at least p_X words of length J leading to a final state. If we order them by genealogical ordering, we denote the p_X first of them by

$$b_i = b_{i,0} < b_{i,1} < \cdots < b_{i,p_X-1}.$$

Notice that for $t \in \{0, \ldots, p_X - 1\}$, we have

$$\mathrm{val}_S(a_i b_{i,t}) = \mathrm{val}_S(a_i b_i) + t = \mathbf{v}_i(q_{0,L}) + t.$$

The sequence $(\mathbf{v}_i(q_{0,L}) \mod p_X)_{i \geq 0}$ is ultimately periodic and takes infinitely often $N_{\mathbf{v}}(p_X) =: N$ different values. Let $h_1, \ldots, h_N \geq J - 1$ such that

$$i \neq j \Rightarrow \mathbf{v}_{h_i}(q_{0,L}) \not\equiv \mathbf{v}_{h_j}(q_{0,L}) \mod p_X$$

and for all $i \in \{1, \ldots, N\}$, $\mathbf{v}_{h_i}(q_{0,L}) \geq a_X$. We have

$$\mathrm{rep}_S(\mathbf{v}_{h_i}(q_{0,L})) = w_{h_i} = a_{h_i} b_{h_i} \text{ and } q_{h_i} = \delta_L(q_{0,L}, a_{h_i}).$$

The elements in the set $\{q_{h_1}, \ldots, q_{h_N}\}$ can take only $\#Q_L$ different values. So at least $\sigma := \lceil N/\#Q_L \rceil$ of them are the same. For the sake of simplicity, assume that they are $q_{h_1}, \ldots, q_{h_\sigma}$. Consequently, for $i, j \in \{1, \ldots, \sigma\}$ and for all $t = 0, \ldots, p_X-1$, we have $b_{h_i,t} = b_{h_j,t}$. For all $i, j \in \{1, \ldots, \sigma\}$ such that $i \neq j$, by Lemma 11, there exists $t_{i,j} < p_X$ such that either $\mathbf{v}_{h_i}(q_{0,L}) + t_{i,j} \in X$ and $\mathbf{v}_{h_j}(q_{0,L}) + t_{i,j} \notin X$ or, $\mathbf{v}_{h_i}(q_{0,L}) + t_{i,j} \notin X$ and $\mathbf{v}_{h_j}(q_{0,L}) + t_{i,j} \in X$. Therefore, the words a_{h_i} and a_{h_j} do not belong to the same equivalence class for the relation $\sim_{\mathrm{rep}_S(X)}$. This can be shown by concatenating the word $b_{h_i,t_{i,j}} = b_{h_j,t_{i,j}}$. Hence the minimal automaton of $\mathrm{rep}_S(X)$ has at least σ states. $\square$

Corollary 40. *Let $S = (L, \Sigma, <)$ be an abstract numeration system having the same properties as in Proposition 39. Assume that the sequence $\mathbf{v} = (\mathbf{v}_i(q_{0,L}))_{i \geq 0}$ is such that*

$$\lim_{m \to +\infty} N_{\mathbf{v}}(m) = +\infty.$$

Then the period of an ultimately periodic set $X \subseteq \mathbb{N}$ such that $\mathrm{rep}_S(X)$ is accepted by a DFA with d states is bounded by the smallest integer s_0 such that for all $m \geq s_0$, $N_{\mathbf{v}}(m) > d \#Q_L$, where Q_L is the set of states of the (trim) minimal automaton of L.

Proposition 41. *Let $S = (L, \Sigma, <)$ be an abstract numeration system. If $X \subseteq \mathbb{N}$ is an ultimately periodic set of period p_X such that $\mathrm{rep}_S(X)$ is accepted by a DFA with d states, then the preperiod a_X of X is bounded by an effectively computable constant C depending only on d and p_X.*

Proof. Let $\mathcal{A} = (Q, q_0, \Sigma, \delta, F)$ be a DFA with d states accepting $\mathrm{rep}_S(X)$. As usual, $\mathcal{M}_L = (Q_L, q_{0,L}, \Sigma, \delta_L, F_L)$ is the minimal automaton of L and for any state $q \in Q_L$, $\mathbf{u}_i(q)$ is the number of words of length i accepted from q in $\mathcal{M}_L$. Since $(\mathbf{u}_i(q))_{i \geq 0}$ satisfies a linear recurrence relation, the sequences $(\mathbf{u}_i(q) \mod p_X)_{i \geq 0}$ are ultimately periodic for all $q \in Q_L$. As usual, we denote by $\iota_{\mathbf{u}(q)}(p_X)$ (resp. $\pi_{\mathbf{u}(q)}(p_X)$) the preperiod (resp. the period) of $(\mathbf{u}_i(q) \mod p_X)_{i \geq 0}$. We set

$$I(p_X) := \max_{q \in Q_L} \iota_{\mathbf{u}(q)}(p_X)$$

and

$$P(p_X) := \mathrm{lcm}_{q \in Q_L} \pi_{\mathbf{u}(q)}(p_X).$$

For a_X large enough, we have $|\operatorname{rep}_S(a_X - 1)| > d \,\#Q_L$. By the pumping lemma applied to the product automaton[3] $\mathcal{A} \times \mathcal{M}_L$, there exist x, y, z with $y \neq \varepsilon$, $|xy| \leq d \,\#Q_L$, $\delta(q_0, x) = \delta(q_0, xy)$, $\delta_L(q_{0,L}, x) = \delta_L(q_{0,L}, xy)$ and such that

$$\operatorname{rep}_S(a_X - 1) = xyz$$

and for all $n \geq 0$,

$$(10) \qquad\qquad\qquad xy^n z \in \operatorname{rep}_S(X).$$

Since $|xy|$ is bounded by a constant, we also have $|z| > I(p_X)$ if a_X is chosen large enough.

Since $|z| > I(p_X)$, using (8), (9) and for all $q \in Q_L$ the periodicity of the sequences $(\mathbf{u}_i(q) \mod p_X)_{i \geq 0}$, we have for all $\ell \geq 0$ that

$$(11) \qquad\qquad \operatorname{val}_S(xy^{\ell p_X P(p_X)} yz) \equiv \operatorname{val}_S(xyz) \mod p_X.$$

Let us give some extra details on how we derive identity (11). Assume $x = x_1 \cdots x_r$, $y = y_1 \cdots y_s$ and $z = z_1 \cdots z_t$. For all $n \geq 1$, using (8) for $w = xy^n z$, we get $|w| = r + ns + t$ and

$$
\begin{aligned}
\operatorname{val}_S(xy^n z) \;=\; \sum_{q \in Q_L} &\left(\sum_{i=1}^{r} \beta_{q,i}(w)\mathbf{u}_{|w|-i}(q) \right. \\
&+ \sum_{i=r+1}^{r+s} \beta_{q,i}(w)\mathbf{u}_{|w|-i}(q) + \cdots + \sum_{i=r+(n-1)s+1}^{r+ns} \beta_{q,i}(w)\mathbf{u}_{|w|-i}(q) \\
&+ \left. \sum_{i=r+ns+1}^{r+ns+t} \beta_{q,i}(w)\mathbf{u}_{|w|-i}(q) \right),
\end{aligned}
$$

where the first (resp. second, third) line corresponds, as explained below, to the contribution of x (resp. y^n, z). By definition (9) of the coefficients $\beta_{q,i}(w)$, we know that $\beta_{q,1}(w)$ depends only on x_1 but $\beta_{q,2}(w)$ depends only on x_2 and on $\delta_L(q_{0,L}, x_1)$. Continuing this way, $\beta_{q,r}(w)$ depends only on x_r and on $\delta_L(q_{0,L}, x_1 \cdots x_{r-1})$ and for $1 \leq j \leq s$, $\beta_{q,r+j}(w)$ depends on y_j and on $\delta_L(q_{0,L}, xy_1 \cdots y_{j-1})$. Now $\beta_{q,r+s+1}(w)$ depends only on y_1 and on $\delta_L(q_{0,L}, xy_1 \cdots y_s) = \delta_L(q_{0,L}, xy) = \delta_L(q_{0,L}, x)$. This implies that $\beta_{q,r+s+j}(w) = \beta_{q,r+j}(w)$ for all $q \in Q_L$ and all $j \in \{1, \ldots, s\}$. This argument can be repeated with every copy of y appearing in w. Consequently, the previous expansion becomes

$$
\begin{aligned}
\operatorname{val}_S(xy^n z) \;=\; \sum_{q \in Q_L} &\left(\sum_{i=1}^{r} \beta_{q,i}(w)\mathbf{u}_{|w|-i}(q) + \sum_{i=r+1}^{r+s} \beta_{q,i}(w) \underbrace{\sum_{j=0}^{n-1} \mathbf{u}_{|w|-i-js}(q)}_{(*)} \right. \\
&+ \left. \sum_{i=r+ns+1}^{r+ns+t} \beta_{q,i}(w)\mathbf{u}_{|w|-i}(q) \right).
\end{aligned}
$$

Assume now that $n = 1 + \ell p_X P(p_X)$, with $\ell \geq 0$. For $q \in Q_L$ and $i = r+1, \ldots, r+s$, we have

$$(*) = \sum_{j=0}^{n-1} \mathbf{u}_{|w|-i-js}(q) = \mathbf{u}_{|w|-i}(q) + \sum_{j=1}^{\ell p_X P(p_X)} \mathbf{u}_{|w|-i-js}(q)$$

---

[3]The automaton $\mathcal{A} \times \mathcal{M}_L$ is defined as follows. For any state (q, q') in the set of states $Q \times Q_L$, when reading $a \in \Sigma$, one reaches in $\mathcal{A} \times \mathcal{M}_L$ the state $(\delta(q, a), \delta_L(q', a))$. The initial state is $(q_0, q_{0,L})$ and the set of final states is $F \times F_L$. Roughly speaking, the product automaton mimics the behavior of both automata $\mathcal{A}$ and $\mathcal{M}_L$.

and the second term is congruent to 0 modulo p_X due to the periodicity of the sequences $(\mathbf{u}_i(q) \mod p_X)_{i \geq 0}$ (recall that in the case we are considering, $|z| = t > I(p_X)$). Consequently, for $n = 1 + \ell p_X P(p_X)$, we have

$$\mathrm{val}_S(xy^n z) \equiv \sum_{q \in Q_L} \left(\sum_{i=1}^{r} \beta_{q,i}(w)\mathbf{u}_{|w|-i}(q) + \sum_{i=r+1}^{r+s} \beta_{q,i}(w)\mathbf{u}_{|w|-i}(q) \right.$$
$$\left. + \sum_{i=r+ns+1}^{r+ns+t} \beta_{q,i}(w)\mathbf{u}_{|w|-i}(q) \right) \mod p_X.$$

It is then easy to derive (11).

We now use the minimality of a_X to get a contradiction. Assume that $a_X - 1$ is in X (the case not in X is similar). Therefore for all $n \geq 1$, $a_X + np_X - 1$ is not in X. From (10), for $\ell > 0$ we get $xy^{\ell p_X P(p_X)}yz \in \mathrm{rep}_S(X)$, but from (11) this word represents a number of the kind $a_X + np_X - 1$ with $n > 0$ which cannot belong to X.

Notice that C can be effectively estimated as follows. One has to choose a constant C such that $a_X > C$ implies $|\mathrm{rep}_S(a_X - 1)| - d \# Q_L > I(p_X)$. Since the abstract numeration system S, the period p_X and the number d of states are given, $I(p_X)$ and $\mathrm{rep}_S(n)$ for all $n \geq 0$ can be effectively computed. □

Theorem 42. *Let $S = (L, \Sigma, <)$ be an abstract numeration system such that for all states q of the trim minimal automaton $\mathcal{M}_L = (Q_L, q_{0,L}, \Sigma, \delta_L, F_L)$ of L*

$$\lim_{i \to \infty} \mathbf{u}_i(q) = +\infty$$

and $\mathbf{u}_i(q_{0,L}) > 0$ for all $i \geq 0$. Assume moreover that $\mathbf{v} = (\mathbf{v}_i(q_{0,L}))_{i \geq 0}$ is such that $\lim_{m \to +\infty} N_{\mathbf{v}}(m) = +\infty$. It is decidable whether or not an S-recognizable set is ultimately periodic.

Proof. The proof is essentially the same as the one of Theorem 22. First notice that the sequence $\mathbf{v} = (\mathbf{v}_i(q_{0,L}))_{i \geq 0}$ satisfies a linear recurrence relation of the kind (1) having a_k as last coefficient. Moreover it is increasing, since $\mathbf{u}_i(q_{0,L}) > 0$ for all $i \geq 0$.

Let the prime decomposition of $|a_k|$ be $|a_k| = p_1^{u_1} \cdots p_r^{u_r}$ with $u_1, \ldots, u_r > 0$. Consider a DFA $\mathcal{A}$ with d states accepting an S-recognizable set $X \subseteq \mathbb{N}$. Assume that X is periodic with period

$$p_X = p_1^{v_1} \cdots p_r^{v_r} c$$

where $\gcd(a_k, c) = 1$ and $v_1, \ldots, v_r \geq 0$.

By Proposition 39, we get $N_{\mathbf{v}}(p_X) \leq d \# Q_L$. Using Remark 17, we obtain

$$N_{\mathbf{v}}(c) \leq \pi_{\mathbf{v}}(c) \leq \pi_{\mathbf{v}}(p_X) \leq (N_{\mathbf{v}}(p_X))^k \leq (d \# Q_L)^k.$$

Let $\alpha(m)$ be defined as the largest index i such that $\mathbf{v}_i(q_{0,L}) < m$. Notice that since the sequence $\mathbf{v}$ is increasing, the map $m \mapsto \alpha(m)$ is non-decreasing and $\lim_{m \to +\infty} \alpha(m) = +\infty$. Since $\gcd(a_k, c) = 1$, the sequence $(\mathbf{v}_i(q_{0,L}) \mod c)_{i \geq 0}$ is purely periodic and $N_{\mathbf{v}}(c) \geq \alpha(c)$. Therefore, $\alpha(c) \leq (d \# Q_L)^k$ and we can give effectively an upper bound on c.

Now we can give upper bound on the v_j's. The assumption $\lim_{m \to +\infty} N_{\mathbf{v}}(m) = +\infty$ implies that $\lim_{v \to +\infty} N_{\mathbf{v}}(p_j^v) = +\infty$ and we have exactly the same reasoning as in the proof of Theorem 22.

We have shown that if X is ultimately periodic, then its period p_X is bounded by a constant that can be effectively estimated. Using Proposition 41, its preperiod is bounded by a constant which can also be computed effectively.

Consequently, the sets of admissible periods and preperiods we have to check are finite. Thanks to Proposition 38, one has to build an automaton for each ultimately periodic set corresponding to a pair of admissible preperiods and periods and then compare the accepted language with $\text{rep}_S(X)$. $\square$

Example 43. The abstract numeration system given in Example 35 satisfies all the assumptions of the previous theorem.

Theorem 42 can be used to decide particular instances of the HD0L periodicity problem. Let Δ, Γ be two finite alphabets. Consider two morphisms $f : \Delta \to \Gamma^*$ and $g : \Delta \to \Delta^*$ such that g is prolongeable on a letter a. The question is to decide whether or not the infinite word $f(g^\omega(a)) = w_0 w_1 w_2 \cdots$ is ultimately periodic. Thanks to [36], one can canonically build an abstract numeration system $S = (L, \Sigma, <)$ and a deterministic finite automaton with output $\mathcal{M} = (Q, q_0, \Sigma, \delta, \Gamma, \tau)$ where $\tau : Q \to \Gamma$ is the output function such that

$$\forall n \geq 0, \ w_n = \tau(\delta(q_0, \text{rep}_S(n))).$$

Such a sequence is said to be an *S-automatic sequence*. Notice that $(w_n)_{n \geq 0}$ is ultimately periodic if and only if for all $b \in \Gamma$, the S-recognizable set

$$X_b = \{n \mid w_n = b\}$$

is ultimately periodic. If f and g are such that the associated numeration system S satisfies the assumptions of Theorem 42, then one can decide whether or not X_b (and therefore $(w_n)_{n \geq 0}$) is ultimately periodic.

Acknowledgments

References

[1] B. Alexeev, Minimal DFAS for testing divisibility, *J. Comput. Syst. Sci.* **69** (2004), 235–243.

[2] J.-P. Allouche, J. Shallit, *Automatic sequences. Theory, applications, generalizations*, Cambridge University Press, Cambridge, (2003).

[3] J.-P. Allouche, N. Rampersad, J. Shallit, Periodicity, repetitions, and orbits of an automatic sequence, to appear in *Theoret. Comput. Sci.*

[4] A. Bertrand-Mathis, Comment écrire les nombres entiers dans une base qui n'est pas entière, *Acta Math. Acad. Sci. Hungar.* **54** (1989), 237–241.

[5] P. B. Bhattacharya, S. K. Jain, S. R. Nagpaul, *Basic abstract algebra*, Second edition, Cambridge University Press, Cambridge, (1994).

[6] J. Bonin, L. Shapiro and R. Simion, Some q-analogues of the Schröder numbers arising from combinatorial statistics on lattice paths, *J. Statist. Plann. Inference* **34** (1993), 35–55.

[7] V. Bruyère, G. Hansel, Bertrand numeration systems and recognizability, Latin American Theoretical INformatics (Valparaíso, 1995), *Theoret. Comput. Sci.* **181** (1997), 17–43.

[8] E. Charlier, M. Rigo, A decision problem for ultimately periodic sets in non-standard numeration systems, Mathematical Foundations of Computer Science 2008 (Toruń), Lecture Notes in Comput. Sci. **5162**, pp. 241–252, Springer-Verlag (2008).

[9] E. Charlier, M. Rigo, W. Steiner, Abstract Numeration Systems on Bounded Languages and Multiplication by a Constant, INTEGERS: Elec. J. of Combin. Number Theory **8(1)** (2008), A35.

[10] F.R.K. Chung and R.L. Graham, On irregularities of distribution of real sequences, *Proc. Nat. Acad. Sci. U.S.A.* **78** (1981), 4001.

[11] F.R.K. Chung and R.L. Graham, On irregularities of distribution, *Finite and Infinite Sets*, Vol. I, II (Eger, 1981), 181–222, Colloq. Math. Soc. János Bolyai, **37** (1984), North-Holland, Amsterdam.

[12] A. Cobham, On the base-dependence of sets of numbers recognizable by finite automata, *Math. Systems Theory* **3** (1969), 186–192.

[13] F. Durand, A theorem of Cobham for non primitive substitution, *Acta Arithmetica* **104** (2002), 225–241.

[14] H. T. Engstrom, On Sequences Defined by Linear Recurrence Relations, *Trans. of the AMS* **33** (1931), 210–218.

[15] G. Everest, A. van der Poorten, I. Shparlinski, T. Ward, *Recurrence sequences*, Mathematical Surveys and Monographs **104**, American Mathematical Society, Providence, RI, (2003).

[16] A.S. Fraenkel, Systems of numeration, *Amer. Math. Monthly* **92** (1985), 105–114.

[17] A.S. Fraenkel, On the recurrence $f_{m+1} = b_m f_m - f_{m-1}$ and applications, *Discrete Mathematics* **224** (2000), 273–279.

[18] A.S. Fraenkel, Arrays, numeration systems and Frankenstein games, *Theoret. Comput. Sci.* **282** (2002), 271–284.

[19] A.S. Fraenkel and S.T. Klein, Novel compression of sparse bit-strings—preliminary report, Proc. NATO Advanced Research Workshop on Combinatorial Algorithms on Words (A. Apostolico and Z. Galil, eds., Maratea, Italy, June 1984), 169–183, Springer-Verlag, Heidelberg, 1985.

[20] C. Frougny, On the sequentiality of the successor function, *Inform. and Computation* **139** (1997), 17–38.

[21] C. Frougny, Representations of numbers and finite automata, *Math. Systems Theory* **25** (1992), 37–60.

[22] C. Frougny, B. Solomyak, On the representation of integers in linear numeration systems, in Ergodic theory of Z_d actions (Warwick, 1993–1994), 345–368, *London Math. Soc. Lecture Note Ser.* **228**, Cambridge Univ. Press, Cambridge, (1996).

[23] F. Q. Gouvêa, *p-adic Numbers: An introduction*, 2nd edition, Springer (1997).

[24] T. Harju, M. Linna, On the periodicity of morphisms on free monoids, *RAIRO Inform. Théor. Appl.* **20** (1986), 47–54.

[25] M. Hollander, Greedy numeration systems and regularity, *Theory Comput. Syst.* **31** (1998), 111–133.

[26] J. Honkala, A decision method for the recognizability of sets defined by number systems, *Theoret. Inform. Appl.* **20** (1986), 395–403.

[27] J. Honkala, M. Rigo, Decidability questions related to abstract numeration systems, *Discrete Math.* **285** (2004), 329–333.

[28] N. Koblitz, *p-adic numbers, p-adic analysis, and zeta-functions*, Springer-Verlag, New York, (1977).

[29] P.B.A. Lecomte, M. Rigo, Numeration systems on a regular language, *Theory Comput. Syst.* **34** (2001), 27–44.

[30] J. Leroux, A Polynomial Time Presburger Criterion and Synthesis for Number Decision Diagrams, 20th IEEE Symposium on Logic in Computer Science (LICS 2005), Chicago, IL, USA, IEEE Computer Society (2005), 147–156.

[31] M. Lothaire, *Algebraic Combinatorics on Words*, Encyclopedia of Mathematics and its Applications **90**, Cambridge University Press, Cambridge, (2002).

[32] A. A. Muchnik, The definable criterion for definability in Presburger arithmetic and its applications, *Theoret. Comput. Sci* **290** (2003) 1433–1444.

[33] J.-J. Pansiot, Decidability of periodicity for infinite words, *RAIRO Inform. Théor. Appl.* **20** (1986), 43–46.

[34] G. Rauzy, Relations de récurrence modulo m, *Séminaire Delange-Pisot-Poitou, Th. Nombres* **5**, (1963-1964).

[35] M. Rigo, Numeration systems on a regular language: arithmetic operations, recognizability and formal power series, *Theoret. Comput. Sci.* **269** (2001), 469–498.

[36] M. Rigo and A. Maes, More on generalized automatic sequences, *J. Autom. Lang. Comb.* **7** (2002), 351–376.

[37] J. Sakarovitch, *Élements de théorie des automates*, Vuibert (2003). English translation: *Elements of Automata Theory*, Cambridge University Press, to appear.

[38] C.D. Savage and A.J. Yee, Euler's partition theorem and the combinatorics of ℓ-sequences, *J. Combinat. Theory, Ser. A* **155** (2008), 967–996.

[39] J. Shallit, Numeration systems, linear recurrences and regular sets, *Inform. and Comput.* **113** (1994) 331–347.

[40] M. Ward, The arithmetical theory of linear recurring series, *Trans. Amer. Math. Soc.* **35** (1933), 600–628.

[41] E. Zeckendorf, Représentation des nombres naturels par une somme des nombres de Fibonacci ou de nombres de Lucas, *Bull. Soc. Roy. Sci. Liège* **41** (1972), 179–182.

(J. Bell) Department of Mathematics, Simon Fraser University Burnaby, BC, CANADA V5A 1S6

E-mail address: jpb@sfu.ca

(E. Charlier, M. Rigo) Institute of Mathematics, University of Liège, Grande Traverse 12 (B 37), B-4000 Liège, Belgium

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

(A. S. Fraenkel) Department of Computer Science & Applied Mathematics, Weizmann Institute of Science, 76100 Rehovot, Israel.

E-mail address: aviezri.fraenkel@weizmann.ac.il