On the regularity of greedy dominant root numeration systems

Émilie Charlier Savinien Kreczman

March 14, 2023

Abstract

In this talk, we build upon the work of Hollander [2] and uncover new criteria allowing us to decide when the set of representations of natural numbers in a greedy numeration system is a regular language.

1 Introduction and definitions

A sequence $(U_n)_{n\in\mathbb{N}}$ satisfying the conditions

$$(i)\ U(0)=1 \quad (ii)\ \forall n\in\mathbb{N},\ U(n+1)>U(n) \quad (iii)\ \exists C\in\mathbb{N},\ \forall n\in\mathbb{N},\ \frac{U_{n+1}}{U_n}\le C$$

can be used to build a numeration system as follows. Let n be a natural number to be represented and $l\in\mathbb{N}$ be such that $U_l\le n<U_{l+1}$, or $l=-1$ if $n=0$. Set $r_l=n$, and then, if r_i is defined and $i\ge 0$, set

$$a_i=\left\lfloor\frac{r_i}{U_i}\right\rfloor \text{ and } r_{i-1}=r_i-a_iU_i.$$

Then the function $\mathrm{rep}\colon\mathbb{N}\to\{0,\dots,C-1\}^*,\ n\mapsto a_l\cdots a_0$, with C chosen minimal, is the *representation function* in the numeration system associated with the sequence U, and it satisfies

$$\mathrm{rep}(n)=a_l\cdots a_0 \implies n=\sum_{i=0}^{l}a_iU_i.$$

This process has been extensively studied (see [1], section 2.2.3 for an overview), with some particular choices of U corresponding to more well-known numeration systems: $U_n=b^n$ corresponds to the usual base b numeration system, and the case where U is the Fibonacci sequence corresponds to a numeration system known as the Zeckendorf numeration system. One can study the *language* of the numeration system, defined by $L_U=\{\mathrm{rep}(n):n\in\mathbb{N}\}$. This language is regular for the two examples above, but it is not when we consider the sequence $U_n=n^2$, as proven by Shallit [5]. One can then ask whether we can characterize the sequences U for which the language L_U is regular.

2 The work of Hollander

This question was introduced by Shallit [5], then studied by Hollander [2]. Shallit proved that U must be a linear recurrence sequence for L_U to be regular. A simpler proof was given by Loraud [3] using tools from formal language theory.

Hollander's work introduced a different set of tools. He first reduced the problem to proving the regularity of the set

$$\operatorname{Maxlg} L_U = \{w \in L_U : v \in L_U,\ |v| = |w| \implies v \preccurlyeq w\}$$

where $\preccurlyeq$ is the lexicographic order. This simplifies the reasoning, as $\operatorname{Maxlg} L_U$ is a language with exactly one word of each length, allowing us to use specific criteria for regularity.

In the case where $\lim_{n\to\infty} \frac{U_{n+1}}{U_n} = \beta > 1$ (this is called the *dominant root condition* for β), Hollander also made the link between the words in this set and the different representations of 1 in the β-numeration system first studied by Rényi [4]. In this numeration system, a number α in $[0,1]$ is represented by the word $a_1 a_2 \cdots$ obtained by the following algorithm. First, set $r_0 = \alpha$. Then, if r_i is defined, let $a_{i+1} = \lfloor \beta r_i \rfloor$ and $r_{i+1} = \beta r_i - a_{i+1}$. Hollander showed that the words in $\operatorname{Maxlg} L_U$ share an arbitrarily long common prefix with the representation of 1 (up to some modification if it ends with a tail of zeroes).

The final tools introduced by Hollander are the β-polynomials, a set of polynomials (one *canonical* and some *extended*) derived from the representation of 1 in the β-numeration system, and an operator associated with a polynomial p: if $p(x) = \sum_{i=0}^{d} c_i x^i$, then

$$\Delta_p \colon \mathbb{Z}^{\mathbb{N}} \to \mathbb{Z}^{\mathbb{N}},\ (U_n)_{n\in\mathbb{N}} \mapsto \left(\sum_{i=0}^{d} c_i U_{n+i}\right)_{n\in\mathbb{N}}.$$

In his article, Hollander proves the following theorem:

Theorem 2.1. *Let U be a linear recurrence sequence satisfying the dominant root condition for $\beta > 1$ and $d_\beta(1)$ be the representation of 1 in the β-numeration system.*

- *If $d_\beta(1)$ is neither finite nor eventually periodic, then L_U is not regular.*
- *If $d_\beta(1)$ is eventually periodic, then L_U is regular if and only if U satisfies an extended β-polynomial.*
- *If $d_\beta(1)$ is finite and U satisfies an extended β-polynomial, then L_U is regular.*
- *If $d_\beta(1)$ is finite and L_U is regular, then U satisfies a polynomial of the form $(x^l - 1)B(x)$ with $B(x)$ an extended β-polynomial and l the length of $d_\beta(1)$.*

In the framework explored by Hollander, with a dominant root greater than 1, only one case remains incompletely characterized: the case where $d_\beta(1)$ is finite of length l and U satisfies some $B(x)(x^l - 1)$ but no extended β-polynomial. In this case, the initial conditions of the sequence U play a role, in addition to the characteristic polynomial of the recurrence relation satisfied by U.

3 New developments

In 2022, we have worked to understand the latter case. Our main tool was to study the precise link between the values of $\Delta_b(U)$ (where $b(x)$ is the canonical β-polynomial) and the remainders seen when applying the greedy algorithm to decompose $U_n - k$. One can prove that, under suitable hypotheses, the remainder after l steps (with l the length of $d_\beta(1)$) when performing the greedy algorithm on $U_n - k$ is $\Delta_b(U)_{n-l} + k$ if this quantity is nonnegative, and $U_{n-l} - k - \Delta_b(U)_{n-l}$ otherwise.

This allows us to prove the following theorem:

Theorem 3.1. *Let U be a linear recurrence sequence with dominant root $\beta > 1$ and $d_\beta(1) = d_1 \cdots d_l$. Suppose that U satisfies the polynomial $x^n(x^{kl} - 1)b(x)$ with b the canonical β-polynomial, $n \in \mathbb{N}$ and $k \in \mathbb{N}_0$. Then the sequence $\Delta_b(U)$ is eventually periodic with period kl. Call this period $(\delta_0, \ldots, \delta_{kl-1})$. Then L_U is regular if, and only if,*

$$\forall j \in \{0, \ldots, l-1\}, \sum_{i=0}^{k-1} \delta_{j+il} \geq 0.$$

This criterion allows us to reprove the third bullet point in the statement of Hollander's theorem, but also fills in the "blank" between the third and fourth points in that statement. The techniques used also give perspective on the difference between the cases where $d_\beta(1)$ is finite or eventually periodic.

We obtain multiple corollaries of this criterion. First, with the hypotheses and notation of Hollander, we prove that a sequence satisfying $(1 + x + \ldots + x^{l-1})B(x)$ is regular if and only if it also satisfies $B(x)$. Second, we have a better understanding of which initial conditions give rise to a regular language L_U, for a given polynomial. We prove that, if U satisfies the polynomial $P(x)B(x)$, with $x - 1$ dividing P and P dividing $x^l - 1$, then the $\deg P$ last initial conditions must be chosen inside an intersection of closed half-spaces in order for L_U to be regular, those half-spaces can be computed effectively and they all pass through a common point.

References

[1] V. Berthé and M. Rigo, editors. *Combinatorics, Automata and Number Theory.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 2010.

[2] M. I. Hollander. Greedy numeration systems and regularity. *Theory Comput. Syst.*, 31(2):111–133, 1998.

[3] N. Loraud. Beta-shift, systèmes de numération et automates. *Journal de Théorie des Nombres de Bordeaux*, 7(2):473–498, 1995. Publisher: Société Arithmétique de Bordeaux.

[4] A. Rényi. Representations for real numbers and their ergodic properties. *Acta Mathematica Academiae Scientiarum Hungaricae*, 8(3-4):477–493, Sept. 1957.

[5] J. Shallit. Numeration Systems, Linear Recurrences, and Regular Sets. *Information and Computation*, 113(2):331–347, Sept. 1994.