# Regularity of positional numeration systems without a dominant root

## Rényi numeration systems

Numeration system: pair of maps, *representation* and *evaluation*, between a set of numbers and a set of words.
Rényi numeration system: between $[0,1]$ and $A^{\mathbb{N}_0}$ for some finite alphabet $A$.
Given a *base* $\beta > 1$, represent a number $x$ using a greedy algorithm. Define $r_0 = x$ then $d_{i+1} = \lfloor \beta r_i \rfloor$ and $r_{i+1} = \beta r_i - d_{i+1}$ if $r_i$ is defined. Obtain a word $d_1 d_2 \cdots$ such that $\sum_{i=1}^{\infty} \frac{d_i}{\beta^i} = x$. This word is called $d_\beta(x)$.

### Example

If $\beta = 10$, we find the usual decimal numeration system.
If $\beta$ is the golden ratio, we have for instance $d_\beta(1/2) = (010)^\omega$ and $d_\beta(1) = 110^\omega$.

## Positional numeration systems

A numeration system between $\mathbb{N}$ and $A^*$ for some finite alphabet $A$. Given an increasing base sequence $(U_n)_{n\in\mathbb{N}}$ with $U_0 = 1$, represent a number $n$ using a greedy algorithm. Let $\ell$ be such that $U_\ell \leq n < U_{\ell+1}$, then set $r_\ell = n$, $a_i = \left\lfloor \frac{r_i}{U_i} \right\rfloor$ and $r_{i-1} = r_i - a_i U_i$. Obtain a word $a_\ell \cdots a_0$ such that $\sum_{i=0}^{\ell} a_i U_i = n$. This word is called $\mathrm{rep}_U(n)$.
The *language* of the numeration system $U$ is $L_U = \{\mathrm{rep}_U(n) : n \in \mathbb{N}\}$. When is this language regular?

### Example

If $U$ is the Fibonacci sequence, we have for instance $\mathrm{rep}_U(20) = 101010$ and $\mathrm{rep}_U(21) = 1000000$. We have $L_U = \{\varepsilon\} \cup 1\{0,1\}^* \setminus \{0,1\}^* 11 \{0,1\}^*$.

When there exists $\beta > 1$ such that $\lim_{n\to\infty} \frac{U_n}{U_{n-1}} = \beta$, Hollander used a link between the $U$- and $\beta$-numeration systems to determine conditions for the regularity of $L_U$. But what if such a $\beta$ does not exist?

## Alternate base numeration systems
## The quasi-greedy algorithm

Alternate base numeration systems extend Rényi numeration systems by involving multiple bases. Consider a sequence of bases $\mathcal{B} = (\beta_n)_{n\in\mathbb{N}}$ that is periodic with period $p$. Numbers in $[0,1]$ are represented with a modified greedy algorithm. Define $r_0 = x$ then $d_{i+1} = \lfloor \beta_i r_i \rfloor$ and $r_{i+1} = \beta_i r_i - d_{i+1}$. Obtain a word $d_1 d_2 \cdots$ such that $\sum_{i=1}^{\infty} \frac{d_i}{\prod_{j=0}^{i-1} \beta_j} = x$. This word is called $d_{\mathcal{B}}(x)$.
Representations in cyclic shifts of the base are also considered. We let $\sigma(\mathcal{B}) = (\beta_{n+1})_{n\in\mathbb{N}}$.
In addition to the representation $d_{\mathcal{B}}(1)$, another representation of $1$ is defined by repeatedly subtracting $1$ from its last digit and appending a new representation of $1$. If $x$ is a word, we let $x^-$ be obtained by subtracting $1$ from the last (nonzero) letter of $x$. Then, we set $w_{i,0} = d_{\sigma^i(\mathcal{B})}(1)$ and, if $w_{i,j}$ is finite, $w_{i,j+1} = (w_{i,j})^- d_{\sigma^{i+|w_{i,j}|}(\mathcal{B})}(1)$. All words defined in this fashion evaluate to $1$ in our numeration system, and they converge to a word noted $d^*_{\sigma^i(\mathcal{B})}(1)$, the *quasi-greedy* representation of $1$.

## Link between $U$- and $\mathcal{B}$-systems

The introduction of alternate base numeration systems is justified because they have a strong link to $U$-systems where $\frac{U_{n+1}}{U_n}$ does not converge.

### Proposition

Let $(U_n)_{n\in\mathbb{N}}$ be a sequence such that $L_U$ is regular. Then there exists $p$ such that the $p$ limits $\lim_{n\to\infty} \frac{U_{pn-i}}{U_{pn-i-1}}$ $(i = 0, \ldots, p-1)$ exist.

We may then construct an alternate base $\mathcal{B}$ by setting $\beta_i = \lim_{n\to\infty} \frac{U_{pn-i}}{U_{pn-i-1}}$ $(i = 0, \ldots, p-1)$ and extending it by periodicity. This base is linked to the initial $U$-system by a result similar to that of Hollander.

### Proposition

For a $U$-system with an associated alternate base $\mathcal{B}$ and for all lengths $\ell$, for all sufficiently large $n$ the word $\mathrm{rep}_U(U_{np-i} - 1)$ shares a prefix of length $\ell$ with one of the words $w_{i,j}$ defined above.

## Example

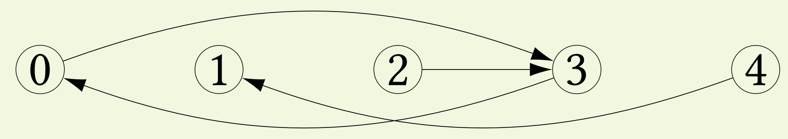
Let $U$ satisfy the recurrence relation $U_{n+10} = 16U_{n+5} - 9U_n$, with the initial conditions below:

| $n$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| $U_n$ | 1 | 2 | 3 | 6 | 10 | 19 | 29 | 48 | 96 | 151 |

We obtain the following values for the $\beta_i$, $d_{\sigma^i(\mathcal{B})}(1)$ and the words $w_{i,j}$.

| $i$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $\beta_i$ | $\frac{11+3\sqrt{55}}{17}$ | $\frac{2+\sqrt{55}}{6}$ | $2$ | $\frac{6+3\sqrt{55}}{17}$ | $\frac{11+3\sqrt{55}}{22}$ |
| $d_{\sigma^i(\mathcal{B})}(1)$ | $1110^\omega$ | $11000(10000)^\omega$ | $20^\omega$ | $110^\omega$ | $110^\omega$ |
| $w_{i,1}$ | $11\mathbf{0}110^\omega$ | / | $\mathbf{1}110^\omega$ | $1\mathbf{0}1110^\omega$ | $1\mathbf{0}11000(10000)^\omega$ |
| $w_{i,2}$ | $1101\mathbf{0}1110^\omega$ | / | $11\mathbf{0}1110^\omega$ | $1011\mathbf{0}110^\omega$ | / |
| $d^*_{\sigma^i(\mathcal{B})}(1)$ | $(11010)^\omega$ | $11000(10000)^\omega$ | $1(10110)^\omega$ | $(10110)^\omega$ | $1011000(10000)^\omega$ |

The following graph $G$ summarizes the behavior of the various classes modulo $p$ through the quasi-greedy algorithm.




## Approach

To decide if $L_U$ is regular, it is enough to decide the regularity of the language $\mathrm{Maxlg}\, L_U$ obtained by extracting the lexicographically largest word of each length from $L_U$. We have $\mathrm{Maxlg}\, L_U = \{\mathrm{rep}_U(U_n - 1) : n \in \mathbb{N}\}$. This language is slender and we can use the above propositions, making it easier to deal with than $L_U$.
In turn, we can split $\mathrm{Maxlg}\, L_U$ into $p$ languages collecting words of different classes modulo $p$ and study their regularity independently. Set $L_i = \{\mathrm{rep}_U(U_{pn-i} - 1) : n \in \mathbb{N}\}$. We study the regularity of the different $L_i$ in order, starting with loops and sinks in the graph $G$ and iteratively moving to their predecessors. This allows us to assume the regularity of successor languages when studying a particular $L_i$.
There are nodes of four types in the graph $G$, and each type has a dedicated criterion deciding the regularity of the language $L_i$.

## Sequences $\Delta$

Our main tool in this work is the study of sequences associated to $U$ which measure "how closely" $U$ matches the representations $d_{\sigma^i(\mathcal{B})}(1)$. When $d_{\sigma^i(\mathcal{B})}(1)$ is finite and equal to $t_1 t_2 \cdots t_\ell$, we set $(\Delta_i)_n = U_{np-i} - \sum_{j=1}^{\ell} t_j U_{np-i-j}$ for all sufficiently large $n$.
Knowledge of these sequences allows us to more accurately deduce the behavior of the $\mathrm{rep}_U(U_{np-i} - 1)$, for instance using the following lemma, which we only state informally to avoid cumbersome notation.

### Lemma (sketch)

Consider $i$ such that $d_{\sigma^j(\mathcal{B})}(1)$ is finite for all $j$ successors of $i$ in the graph. We know that $\mathrm{rep}_U(U_{np-i} - 1)$ shares a long common prefix with some $w_{i,j}$. There exists a sequence of indices along which to extract values of $\Delta$ such that, if the cumulative sums of the first $j_0$ entries are nonpositive and the next cumulative sum is positive, then the common prefix mentioned above is with $w_{i,j_0}$. In this case, we also have information on $d_{\sigma^i(\mathcal{B})}(1)$ after it diverges from $w_{i,j_0}$.

## Main results

We obtain criteria linking the regularity of $L_i$ to the behavior of $\Delta_i$.

### Theorem

Consider $i$ such that it is not a sink in the graph but there is a path from $i$ to a sink. Assume that the languages $L_j$ are regular for all successors $j$ of $i$ in the graph.
Then, $L_i$ is regular if, and only if, $\Delta_i$ is ultimately periodic.

There are three other theorems relating to the other three cases, but we do not present them as they require additional notation and are more technical.

### Example

In the example on top of this column, all sequences $\Delta$ are identically zero, so the language $L_U$ is regular.

## Reference


[1] M. I. Hollander, *Greedy numeration systems and regularity*, Theory Comput. Syst., vol. 31, n° 2, p. 111-133, 1998.



Émilie Charlier, Savinien Kreczman
{echarlier,savinien.kreczman}@uliege.be


LIÈGE université
fnrs LA LIBERTÉ DE CHERCHER