

The regularity of languages generated by a greedy numeration system

Savinien Kreczman
Joint work with Émilie Charlier

June 5, 2026



1 Introduction

2 Hollander

3 Beyond

U -systems

U -systems: *base* sequence (U_n) , the word $w_{\ell-1} \cdots w_0$ has value $\sum_{j=0}^{\ell-1} w_j U_j$.

Example

The *Zeckendorf* numeration system is based on the Fibonacci sequence.

n	55	34	21	13	8	5	3	2	1
42	0	1	0	0	1	0	0	0	0
42	0	0	1	1	0	1	0	1	1
42	0	0	2	0	0	0	0	0	0

U -systems

U -systems: *base sequence* (U_n) , the word $w_{\ell-1} \cdots w_0$ has value $\sum_{j=0}^{\ell-1} w_j U_j$.

We use a *greedy algorithm* to represent numbers. Let ℓ be such that $U_{\ell-1} \leq n < U_\ell$ and $r_\ell = n$, then set $a_i = \left\lfloor \frac{r_{i+1}}{U_i} \right\rfloor$ and $r_i = r_{i+1} - a_i U_i$

Example

The *Zeckendorf* numeration system is based on the Fibonacci sequence.

n	55	34	21	13	8	5	3	2	1
$\rightarrow 42$	0	1	0	0	1	0	0	0	0
42	0	0	1	1	0	1	0	1	1
42	0	0	2	0	0	0	0	0	0

The *language* of all representations is the set
 $L_U = 0^* \{\text{rep}_U(n) : n \in \mathbb{N}\}.$

The *language* of all representations is the set

$$L_U = 0^* \{\text{rep}_U(n) : n \in \mathbb{N}\}.$$

Decimal numeration system : $L_U = \{0, \dots, 9\}^*$

Zeckendorf numeration system : $L_U = \{0, 1\}^* \setminus \{0, 1\}^* 11 \{0, 1\}^*$

Regularity

The *language* of all representations is the set

$$L_U = 0^* \{\text{rep}_U(n) : n \in \mathbb{N}\}.$$

Decimal numeration system : $L_U = \{0, \dots, 9\}^*$

Zeckendorf numeration system : $L_U = \{0, 1\}^* \setminus \{0, 1\}^* 11 \{0, 1\}^*$

Squares numeration system : L_U is not regular.

U_j	$\dots$	$(j+1)^2$	j^2	$\dots$	$\sim 2j+1$	$\dots$
$\text{rep}_U(n)$	$\dots$	0	1	0...0	*	$\dots$

Regularity

The *language* of all representations is the set

$$L_U = 0^* \{\text{rep}_U(n) : n \in \mathbb{N}\}.$$

Decimal numeration system : $L_U = \{0, \dots, 9\}^*$

Zeckendorf numeration system : $L_U = \{0, 1\}^* \setminus \{0, 1\}^* 11 \{0, 1\}^*$

Squares numeration system : L_U is not regular.

$$\begin{array}{c|ccccccc} U_j & \cdots & (j+1)^2 & j^2 & \cdots & \sim 2j+1 & \cdots \\ \hline \text{rep}_U(n) & \cdots & 0 & 1 & 0 \cdots 0 & * & \cdots \end{array}$$

Question

Can we characterize the sequences U such that L_U is regular?

Proposition (Shallit 1994)

If L_U is regular, then U is a linear recurrence sequence.

Proposition (Hollander 1998)

L_U is regular if and only if

$$\text{Max}(L_U) := \{\text{rep}_U(U_n - 1) : n \in \mathbb{N}\}$$

is regular.

For the Zeckendorf numeration system:

$$\text{Max}(L_U) = \{\varepsilon, 1, 10, 101, 1010, 10101, \dots\} = \text{Pref}((10)^\omega)$$

Proposition (Shallit 1994)

If L_U is regular, then U is a linear recurrence sequence.

Proposition (Hollander 1998)

L_U is regular if and only if

$$\text{Max}(L_U) := \{\text{rep}_U(U_n - 1) : n \in \mathbb{N}\}$$

is regular.

For the Zeckendorf numeration system:

$$\text{Max}(L_U) = \{\varepsilon, 1, 10, 101, 1010, 10101, \dots\} = \text{Pref}((10)^\omega)$$

Rényi numeration systems use a base $\beta > 1$ with the evaluation map

$$\text{val}: a_{-1}a_{-2}\cdots \mapsto \sum_{j=1}^{\infty} a_{-j}\beta^{-j}$$

and the greedy algorithm defined as follows. Set $r_0 = x$ then set $a_i = \lfloor \beta r_{i+1} \rfloor$ and $r_i = \beta r_{i+1} - a_i = T_\beta(r_{i+1})$. We have $\text{rep}_\beta(x) = a_{-1}a_{-2}\cdots$, which we call $d_\beta(x)$.

Example

Let us take β to be the golden ratio φ . We have $\text{rep}(2/\varphi^2) = 1001$.

Let us consider representations of 1. In base φ , the representation of 1 is 11, which means the following words have value 1:

$$w_0 = 11, w_1 = 1011, w_2 = 101011, \dots, w_\infty = (10)^\omega.$$

Hollander's connection

Let us consider representations of 1. In base φ , the representation of 1 is 11, which means the following words have value 1:

$$w_0 = 11, w_1 = 1011, w_2 = 101011, \dots, w_\infty = (10)^\omega.$$

Hollander's connection

Let us consider representations of 1. In base φ , the representation of 1 is 11, which means the following words have value 1:

$$w_0 = 11, w_1 = 1011, w_2 = 101011, \dots, w_\infty = (10)^\omega.$$

This is a strong instance of a result by Hollander.

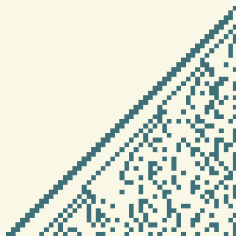
Theorem (Hollander, 1998)

If the numeration system U is such that $\frac{U_{n+1}}{U_n} \rightarrow \beta > 1$ then, given ℓ , for n large enough $\text{rep}_U(U_n - 1)$ shares a prefix of length at least ℓ with some w_i .

We say that U has a *dominant root* β .

A worse case

An example of a U where Hollander's criterion applies less strongly is the sequence U given by the recurrence relation $U_{n+4} = 2U_{n+3} + U_{n+2} - 2U_{n+1} - U_n$ and the initial conditions $(1, 2, 3, 6)$. This system is linked to the golden ratio.



This image represents the first 50 words in $\text{Max}(L_U)$ (which have length 0 to 49). Each line represents a word, each teal square a digit 1 and each cream square a digit 0.

Before finally stating Hollander's result, we must define β -polynomials, which will help us control the greedy algorithm.

Definition

If $d_\beta(1) = t_1 t_2 \dots t_\ell$, the β -polynomial is $X^\ell - \sum_{i=1}^{\ell} t_i X^{\ell-i}$.

If $d_\beta(1) = t_1 t_2 \dots t_q (t_{q+1} \dots t_{q+\ell})^\omega$, the β -polynomial is

$$X^{q+\ell} - \sum_{i=1}^{q+\ell} t_i X^{q+\ell-i} - \left(X^q - \sum_{i=1}^q t_i X^{q-i} \right)$$

The *extended β -polynomials* are obtained by multiplying these by any $X^m(1 + X^\ell + \dots + X^{(a-1)\ell})$.

Theorem (Hollander 1998)

$d_\beta(1)$ infinite periodic	$d_\beta(1) = d_1 \cdots d_l$	$d_\beta(1)$ infinite aperiodic
U doesn't follow an extended β -polynomial	U doesn't follow an extended β -polynomial multiplied by a divisor of $x^\ell - 1$	
	U follows an extended β -polynomial multiplied by a divisor of $x^\ell - 1$	
U follows an extended β -polynomial	U follows an extended β -polynomial	

 Regular language
  Non-regular language

Not every numeration system with a regular language has a dominant root!

Example

Consider the sequence $U_n = (1, 10, 60, 600, 3600, 36000, \dots)$.
 L_U is regular but $\frac{U_{n+1}}{U_n}$ alternates between 10 and 6.

Not every numeration system with a regular language has a dominant root!

Example

Consider the sequence $U_n = (1, 10, 60, 600, 3600, 36000, \dots)$.
 L_U is regular but $\frac{U_{n+1}}{U_n}$ alternates between 10 and 6.

We introduce *alternate bases* as our answer. For instance, the system above would be associated with the alternate base $(6, 10)$

Alternate bases

An alternate base is given by p bases $(\beta_{p-1}, \dots, \beta_0)$, which we extend to $\mathbb{Z}$ by periodicity.

The greedy algorithm alternates between the p transformations

$T_{\beta_{p-1}}, \dots, T_{\beta_0}$.

The evaluation map is given by

$$\text{val}: a_{-1}a_{-2}\cdots \mapsto \sum_{j=1}^{\infty} \frac{a_{-j}}{\beta_{-1}\cdots\beta_{-j}},$$

Alternate bases

An alternate base is given by p bases $(\beta_{p-1}, \dots, \beta_0)$, which we extend to $\mathbb{Z}$ by periodicity.

The greedy algorithm alternates between the p transformations

$$T_{\beta_{p-1}}, \dots, T_{\beta_0}.$$

The evaluation map is given by

$$\text{val}: a_{-1}a_{-2}\cdots \mapsto \sum_{j=1}^{\infty} \frac{a_{-j}}{\beta_{-1}\cdots\beta_{-j}},$$

Example

Let us compute $d_{\mathcal{B}}(1)$ where $\mathcal{B} = \left(\beta_1 = \frac{1+\sqrt{13}}{2}, \beta_0 = \frac{5+\sqrt{13}}{6}\right)$.

i	0	-1	-2	-3
r_i	1			
$\beta_{i-1}r_i$				
a_{i-1}				

Alternate bases

An alternate base is given by p bases $(\beta_{p-1}, \dots, \beta_0)$, which we extend to $\mathbb{Z}$ by periodicity.

The greedy algorithm alternates between the p transformations

$$T_{\beta_{p-1}}, \dots, T_{\beta_0}.$$

The evaluation map is given by

$$\text{val}: a_{-1}a_{-2}\cdots \mapsto \sum_{j=1}^{\infty} \frac{a_{-j}}{\beta_{-1}\cdots\beta_{-j}},$$

Example

Let us compute $d_{\mathcal{B}}(1)$ where $\mathcal{B} = \left(\beta_1 = \frac{1+\sqrt{13}}{2}, \beta_0 = \frac{5+\sqrt{13}}{6}\right)$.

i	0	-1	-2	-3
r_i	1			
$\beta_{i-1}r_i$	β_1			
a_{i-1}				

Alternate bases

An alternate base is given by p bases $(\beta_{p-1}, \dots, \beta_0)$, which we extend to $\mathbb{Z}$ by periodicity.

The greedy algorithm alternates between the p transformations

$$T_{\beta_{p-1}}, \dots, T_{\beta_0}.$$

The evaluation map is given by

$$\text{val}: a_{-1}a_{-2}\cdots \mapsto \sum_{j=1}^{\infty} \frac{a_{-j}}{\beta_{-1}\cdots\beta_{-j}},$$

Example

Let us compute $d_{\mathcal{B}}(1)$ where $\mathcal{B} = \left(\beta_1 = \frac{1+\sqrt{13}}{2}, \beta_0 = \frac{5+\sqrt{13}}{6}\right)$.

i	0	-1	-2	-3
r_i	1	$\beta_1 - 2$		
$\beta_{i-1}r_i$	β_1			
a_{i-1}	2			

Alternate bases

An alternate base is given by p bases $(\beta_{p-1}, \dots, \beta_0)$, which we extend to $\mathbb{Z}$ by periodicity.

The greedy algorithm alternates between the p transformations

$$T_{\beta_{p-1}}, \dots, T_{\beta_0}.$$

The evaluation map is given by

$$\text{val}: a_{-1}a_{-2}\cdots \mapsto \sum_{j=1}^{\infty} \frac{a_{-j}}{\beta_{-1}\cdots\beta_{-j}},$$

Example

Let us compute $d_{\mathcal{B}}(1)$ where $\mathcal{B} = \left(\beta_1 = \frac{1+\sqrt{13}}{2}, \beta_0 = \frac{5+\sqrt{13}}{6}\right)$.

i	0	-1	-2	-3
r_i	1	$\beta_1 - 2$	$(\beta_1 - 2)\beta_0$	
$\beta_{i-1}r_i$	β_1	$(\beta_1 - 2)\beta_0$		
a_{i-1}	2	0		

Alternate bases

An alternate base is given by p bases $(\beta_{p-1}, \dots, \beta_0)$, which we extend to $\mathbb{Z}$ by periodicity.

The greedy algorithm alternates between the p transformations

$$T_{\beta_{p-1}}, \dots, T_{\beta_0}.$$

The evaluation map is given by

$$\text{val}: a_{-1}a_{-2}\cdots \mapsto \sum_{j=1}^{\infty} \frac{a_{-j}}{\beta_{-1}\cdots\beta_{-j}},$$

Example

Let us compute $d_{\mathcal{B}}(1)$ where $\mathcal{B} = \left(\beta_1 = \frac{1+\sqrt{13}}{2}, \beta_0 = \frac{5+\sqrt{13}}{6}\right)$.

i	0	-1	-2	-3
r_i	1	$\beta_1 - 2$	$(\beta_1 - 2)\beta_0$	0
$\beta_{i-1}r_i$	β_1	$(\beta_1 - 2)\beta_0$	$(\beta_1 - 2)\beta_0\beta_1$	0
a_{i-1}	2	0	1	0

Alternate bases

An alternate base is given by p bases $(\beta_{p-1}, \dots, \beta_0)$, which we extend to $\mathbb{Z}$ by periodicity.

The greedy algorithm alternates between the p transformations

$$T_{\beta_{p-1}}, \dots, T_{\beta_0}.$$

The evaluation map is given by

$$\text{val}: a_{-1}a_{-2}\cdots \mapsto \sum_{j=1}^{\infty} \frac{a_{-j}}{\beta_{-1}\cdots\beta_{-j}},$$

Example

Let us compute $d_{\mathcal{B}}(1)$ where $\mathcal{B} = \left(\beta_1 = \frac{1+\sqrt{13}}{2}, \beta_0 = \frac{5+\sqrt{13}}{6}\right)$.

i	0	-1	-2	-3
r_i	1	$\beta_1 - 2$	$(\beta_1 - 2)\beta_0$	0
$\beta_{i-1}r_i$	β_1	$(\beta_1 - 2)\beta_0$	$(\beta_1 - 2)\beta_0\beta_1$	0
a_{i-1}	2	0	1	0

So $d_{\mathcal{B}}(1) = 201$.

Alternate bases

An alternate base is given by p bases $(\beta_{p-1}, \dots, \beta_0)$, which we extend to $\mathbb{Z}$ by periodicity.

The greedy algorithm alternates between the p transformations

$$T_{\beta_{p-1}}, \dots, T_{\beta_0}.$$

The evaluation map is given by

$$\text{val}: a_{-1}a_{-2}\cdots \mapsto \sum_{j=1}^{\infty} \frac{a_{-j}}{\beta_{-1}\cdots\beta_{-j}},$$

Example

Let us compute $d_{\mathcal{B}}(1)$ where $\mathcal{B} = \left(\beta_1 = \frac{1+\sqrt{13}}{2}, \beta_0 = \frac{5+\sqrt{13}}{6}\right)$.

i	0	-1	-2	-3
r_i	1	$\beta_1 - 2$	$(\beta_1 - 2)\beta_0$	0
$\beta_{i-1}r_i$	β_1	$(\beta_1 - 2)\beta_0$	$(\beta_1 - 2)\beta_0\beta_1$	0
a_{i-1}	2	0	1	0

So $d_{\mathcal{B}}(1) = 201$. To find $d_{\mathcal{B}}^*(1)$, we must know $d_{\sigma(\mathcal{B})}(1) = 11$.

A suitable choice

Alternate bases really are the correct extensions of Rényi systems when working on the regularity of L_U .

Proposition (Charlier, K. 2025)

If L_U is regular, U is associated to an alternate base: there exist p and $\beta_{p-1}, \dots, \beta_0$ such that $\frac{U_{np+i+1}}{U_{np+i}} \rightarrow \beta_i$ for $i \in \{0, \dots, p-1\}$.

A suitable choice

Alternate bases really are the correct extensions of Rényi systems when working on the regularity of L_U .

Proposition (Charlier, K. 2025)

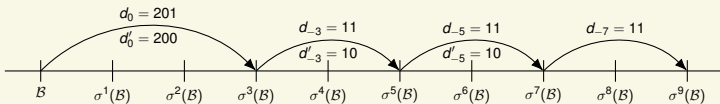
If L_U is regular, U is associated to an alternate base: there exist p and $\beta_{p-1}, \dots, \beta_0$ such that $\frac{U_{np+i+1}}{U_{np+i}} \rightarrow \beta_i$ for $i \in \{0, \dots, p-1\}$.

Hollander's connection generalizes to alternate bases.

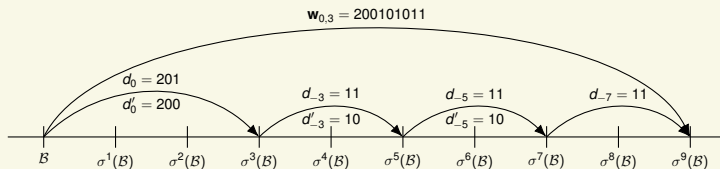
Theorem (Charlier, K. 2025)

If the numeration system U is associated with the alternate base $(\beta_{p-1}, \dots, \beta_0)$ then, given ℓ , for n large enough $\text{rep}_U(U_{np+i} - 1)$ shares a prefix of length at least ℓ with some intermediate representation $\mathbf{w}_{i,j}$ of 1 in the base $(\beta_{i-1}, \beta_{i-2}, \dots, \beta_{i-p})$.

Step 1: sublanguages modulo p

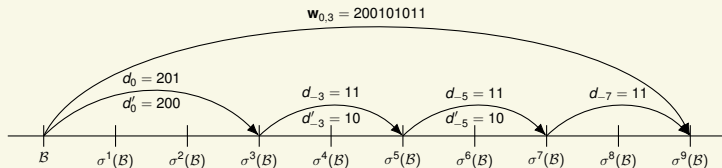


Step 1: sublanguages modulo p



To obtain $w_{i,j}$: start at $i \bmod p$, choose the non-greedy representation j times then the greedy representation.

Step 1: sublanguages modulo p



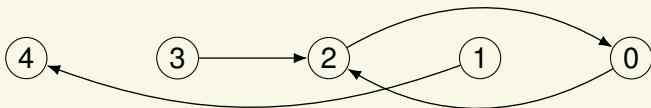
To obtain $\mathbf{w}_{i,j}$: start at $i \bmod p$, choose the non-greedy representation j times then the greedy representation.
Idea: study the regularity of $\text{Max}(L_U) \cap A^i(A^n)^*$ separately for each i , according to its "routing class".

Step 1: sublanguages modulo p - graph for routing

Consider a base $\mathcal{B}$ where the extensions of 1 are

$$\mathbf{d}_{-1} = 11(00010)^\omega, \mathbf{d}_{-2} = 20^\omega, \mathbf{d}_{-3} = \mathbf{d}_{-4} = 110^\omega, \mathbf{d}_{-5} = 1110^\omega,$$

extended by periodicity.



Step 2: Finely controlling the greedy algorithm

Recall the β -polynomials:

$$X^\ell - \sum_{i=1}^{\ell} t_i X^{\ell-i} \text{ if } d_\beta(1) = t_1 \cdots t_\ell$$

$$X^{q+\ell} - \sum_{i=1}^{q+\ell} t_i X^{q+\ell-i} - (X^q - \sum_{i=1}^q t_i X^{q-i}) \text{ if } d_\beta(1) \text{ infinite}$$

Step 2: Finely controlling the greedy algorithm

Recall the β -polynomials:

$$X^\ell - \sum_{i=1}^{\ell} t_i X^{\ell-i} \text{ if } d_\beta(1) = t_1 \cdots t_\ell$$

$$X^{q+\ell} - \sum_{i=1}^{q+\ell} t_i X^{q+\ell-i} - (X^q - \sum_{i=1}^q t_i X^{q-i}) \text{ if } d_\beta(1) \text{ infinite}$$

Definition

If $\mathbf{d}_j$ is finite equal to $t_{i,1} \cdots t_{i,\ell_i}$, we define the sequence (Δ_i) by

$$(\Delta_i)_n = U_{np+i} - \sum_{\ell=1}^{\ell_i} t_{i,\ell} U_{np+i-\ell}.$$

If $\mathbf{d}_j$ is ultimately periodic with period q and preperiod m ,

$$(\Delta_{i,q,m})_n = (U_{np+i} - \sum_{\ell=1}^{q+m} t_{i,\ell} U_{np+i-\ell}) - (U_{np+i-m} - \sum_{\ell=1}^q t_{i,\ell} U_{np+i-m-\ell}).$$

Step 2: Finely controlling the greedy algorithm

Let $(\Delta_i^{(j)})_n$ be the cumulative sum of j terms of Δ starting from n . These sequences enable us to finely control the greedy algorithm.

Proposition (Charlier, K. 2025)

For large enough n ,

- *If $(\Delta_i^{(h)})_n < 1$ for all $h \in \{1, \dots, j\}$, then*

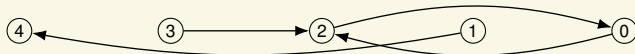
$$\text{rep}_U(U_{np+i} - 1) = d'_i d'_{i-k_{i,1}} \cdots d'_{i-k_{i,j-1}} \text{rep}_U(U_{np+i-k_{i,j}} - c_n),$$

where $c_n = 1 - (\Delta_i^{(j)})_n$;

- *If $(\Delta_i^{(h)})_n < 1$ for all $h \in \{1, \dots, j-1\}$ and $(\Delta_i^{(j)})_n \geq 1$, then*

$$\text{rep}_U(U_{np+i} - 1) \in d'_i \cdots d'_{i-k_{i,j-2}} d_{i-k_{i,j-1}} 0^* \text{rep}_U((\Delta_i^{(j)})_n - 1).$$

Step 3: do the rest of the work



We obtain a characterization of the regularity of $L_i = \text{Max}(L_U) \cap A^i(A^n)^*$ depending on the behavior of i in the routing graph, and assuming that all successors are regular.

Theorem (Charlier, K. 2025 ; sketch)

- *Case 4: L_i is regular iff $\Delta_{i,q,km}$ is eventually 0 for some k .*
- *Case 1: L_i is regular iff Δ_i is eventually periodic.*
- *Case 0/2: If all the L_i in the loop are regular, all the Δ_i are eventually periodic (say of period M). If the $\Delta_i^{(M)}$ are all eventually nonnegative, the L_i are all regular.*
- *Case 3: L_i is regular iff Δ_i is eventually a merge of affine sequences, with a very strict condition on the linear terms.*

- We reobtain Hollander's criteria as byproducts of our own.
- We also complete the grey area in Hollander's results. For a given polynomial where Hollander's criteria are not sufficient, the initial conditions that give a regular language form a polyhedral cone.
- New behaviors can occur, notably in cases 1 and 3.
- Our criteria are effective provided that we know the eigenvalues of U and that the representations of 1 in the associated alternate base are finite or eventually periodic.
- Our results seem to disprove Hollander's original conjecture regarding the alternate case.

Thank you for your attention!