Numeration systems and string attractors

France Gheeraert
Joint work with Giuseppe Romana and Manon Stipulanti

May 26, 2023

String attractors

Definition

A *string attractor* for a (finite) word w is a set of positions $\Gamma \subseteq \{1, \ldots, |w|\}$ covering each factor of w, i.e. each factor has an occurrence in w crossing one of the positions in Γ.

For the word

$$01000101,$$

String attractors

Definition

A *string attractor* for a (finite) word w is a set of positions $\Gamma \subseteq \{1, \ldots, |w|\}$ covering each factor of w, i.e. each factor has an occurrence in w crossing one of the positions in Γ.

For the word

$$010\underline{0}0\underline{1}01,$$

the set $\{4, 6\}$ is a string attractor.

Definition

A *string attractor* for a (finite) word w is a set of positions $\Gamma \subseteq \{1, \dots, |w|\}$ covering each factor of w, i.e. each factor has an occurrence in w crossing one of the positions in Γ.

For the word

$$010\underline{0}0\underline{1}01,$$

length	factors	covered?
1	0, 1	

the set $\{4, 6\}$ is a string attractor.

String attractors

Definition

A *string attractor* for a (finite) word w is a set of positions $\Gamma \subseteq \{1, \ldots, |w|\}$ covering each factor of w, i.e. each factor has an occurrence in w crossing one of the positions in Γ.

For the word

$$010\underline{0}0\underline{1}01,$$

length	factors	covered?
1	0, 1	✓

the set $\{4, 6\}$ is a string attractor.

String attractors

Definition

A *string attractor* for a (finite) word w is a set of positions $\Gamma \subseteq \{1, \dots, |w|\}$ covering each factor of w, i.e. each factor has an occurrence in w crossing one of the positions in Γ.

For the word

$010\underline{0}0\underline{1}01$,

length	factors	covered?
1	0, 1	✓
2	01, 10, 00	

the set $\{4, 6\}$ is a string attractor.

String attractors

Definition

A *string attractor* for a (finite) word w is a set of positions $\Gamma \subseteq \{1, \ldots, |w|\}$ covering each factor of w, i.e. each factor has an occurrence in w crossing one of the positions in Γ.

For the word

$010\underline{0}01\underline{1}01,$

length	factors	covered?
1	0, 1	✓
2	01, 10, 00	

the set $\{4, 6\}$ is a string attractor.

String attractors

Definition

A *string attractor* for a (finite) word w is a set of positions $\Gamma \subseteq \{1, \ldots, |w|\}$ covering each factor of w, i.e. each factor has an occurrence in w crossing one of the positions in Γ.

For the word

$$010\underline{0}0\underline{1}01,$$

length	factors	covered?
1	0, 1	✓
2	01, 10, 00	

the set $\{4, 6\}$ is a string attractor.

String attractors

Definition

A *string attractor* for a (finite) word w is a set of positions $\Gamma \subseteq \{1, \ldots, |w|\}$ covering each factor of w, i.e. each factor has an occurrence in w crossing one of the positions in Γ.

For the word

$01\underline{0}0\underline{1}01$,

length	factors	covered?
1	0, 1	✓
2	01, 10, 00	✓

the set $\{4, 6\}$ is a string attractor.

String attractors

Definition

A *string attractor* for a (finite) word w is a set of positions $\Gamma \subseteq \{1, \ldots, |w|\}$ covering each factor of w, i.e. each factor has an occurrence in w crossing one of the positions in Γ.

For the word

$010\underline{0}0\underline{1}01$,

length	factors	covered?
1	0, 1	✓
2	01, 10, 00	✓
3	010, 100, 000, 001, 101	

the set $\{4, 6\}$ is a string attractor.

String attractors

Definition

A *string attractor* for a (finite) word w is a set of positions $\Gamma \subseteq \{1, \ldots, |w|\}$ covering each factor of w, i.e. each factor has an occurrence in w crossing one of the positions in Γ.

For the word

$010\underline{0}0\underline{1}01$,

length	factors	covered?
1	0, 1	✓
2	01, 10, 00	✓
3	010, 100, 000, 001, 101	✓

the set $\{4, 6\}$ is a string attractor.

String attractors

Definition

A *string attractor* for a (finite) word w is a set of positions $\Gamma \subseteq \{1, \ldots, |w|\}$ covering each factor of w, i.e. each factor has an occurrence in w crossing one of the positions in Γ.

For the word

$0\underline{1}0\underline{00}101$,

length	factors	covered?
1	0, 1	
2	01, 10, 00	
3	010, 100, 000, 001, 101	

the set $\{4, 6\}$ is a string attractor but $\{2, 4, 5\}$ is not.

String attractors

Definition

A *string attractor* for a (finite) word w is a set of positions $\Gamma \subseteq \{1, \ldots, |w|\}$ covering each factor of w, i.e. each factor has an occurrence in w crossing one of the positions in Γ.

For the word

$0\underline{1}0\underline{00}101$,

length	factors	covered?
1	0, 1	✓
2	01, 10, 00	✓
3	010, 100, 000, 001, 101	✗

the set $\{4, 6\}$ is a string attractor but $\{2, 4, 5\}$ is not.

Why string attractors?

- String attractors were introduced by Prezza in 2017 for data compression.

Why string attractors?

- String attractors were introduced by Prezza in 2017 for data compression.
- Finding a string attractor of minimal size for a given word is an NP-complete problem (Kampa-Prezza '18)

Why string attractors?

- String attractors were introduced by Prezza in 2017 for data compression.
- Finding a string attractor of minimal size for a given word is an NP-complete problem (Kampa-Prezza '18)
- That's where combinatorists on words can help! (Mantaci *et al.* '21, Schaeffer-Shallit '20, Restivo-Romana-Sciortino '22, Dvořáková '22, etc.)

Why string attractors?

- String attractors were introduced by Prezza in 2017 for data compression.
- Finding a string attractor of minimal size for a given word is an NP-complete problem (Kampa-Prezza '18)
- That's where combinatorists on words can help! (Mantaci *et al.* '21, Schaeffer-Shallit '20, Restivo-Romana-Sciortino '22, Dvořáková '22, etc.)
- We want to find a string attractor for each prefix of some infinite word.

Two famous infinite words

Definition

A *morphism* (or substitution) is an application respecting the concatenation, i.e., for all words u, v, we have $\mu(uv) = \mu(u)\mu(v)$.

If $\mu(0) = 0u$ with u such that for all $n \geq 1$, $\mu^n(u)$ is not empty, then we define the *fixed point* of μ as

$$\mu^\omega(0) = \lim_{n \to \infty} \mu^n(0).$$

Two famous infinite words

Definition

A *morphism* (or substitution) is an application respecting the concatenation, i.e., for all words u, v, we have $\mu(uv) = \mu(u)\mu(v)$.

If $\mu(0) = 0u$ with u such that for all $n \geq 1$, $\mu^n(u)$ is not empty, then we define the *fixed point* of μ as

$$\mu^\omega(0) = \lim_{n\to\infty} \mu^n(0).$$

Examples

1. μ such that $\mu(0) = 01$, $\mu(1) = 10$
2. μ such that $\mu(0) = 01$, $\mu(1) = 0$

Two famous infinite words

Definition

A *morphism* (or substitution) is an application respecting the concatenation, i.e., for all words u, v, we have $\mu(uv) = \mu(u)\mu(v)$.

If $\mu(0) = 0u$ with u such that for all $n \geq 1$, $\mu^n(u)$ is not empty, then we define the *fixed point* of μ as

$$\mu^\omega(0) = \lim_{n\to\infty} \mu^n(0).$$

Examples

1. μ such that $\mu(0) = 01$, $\mu(1) = 10$: **Thue–Morse word** $\longrightarrow$ string attractors of size 4 (Katsukake *et al.* '20, Schaeffer-Shallit '21)
2. μ such that $\mu(0) = 01$, $\mu(1) = 0$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^{\omega}(0) = \underline{0}100101001001\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = \underline{01}00101001001\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = \underline{01}0\textcolor{gray}{0101001001}\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = \underline{01}00101001001\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = \underline{010}0101001001\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = \underline{010}0101001001\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = 0\underline{10}0101001001\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = 0\underline{10}01\,01001001\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = 0\underline{10}010\textcolor{gray}{1001001}\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = 0\underline{10}0{\color{blue}101}{\color{gray}001001}\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = 0\underline{10}0\underline{1}01001001\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = 0\underline{10}0\underline{1}01001001\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = 01\underline{0}0\underline{1}01001001 \cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = 01\underline{0}0\underline{1}010\textcolor{gray}{01001}\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = 01\underline{0}0\underline{1}0100\textcolor{gray}{1001}\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = 01\underline{0}0\underline{1}01001001\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = 01\underline{0}0\underline{1}01001001\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = 01\underline{0}0\underline{1}01{\color{blue}00100}{\color{gray}1 \cdots}$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^{\omega}(0) = 01\underline{0}0\underline{1}01\underline{0}01001\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^{\omega}(0) = 01\underline{0}0\underline{1}01\underline{0}01001\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = 0100\underline{1}01\underline{0}01001\cdots$$

Fixed point of $\mu(0) = 01$ and $\mu(1) = 0$

Construction of the string attractors:

$$\mu^\omega(0) = 0100\underline{1}01\underline{0}01001\cdots$$

We have the string attractors $\{1,2\}$, $\{2,3\}$, $\{3,5\}$, $\{5,8\}$, ...

$\longrightarrow$ Fibonacci sequence!

Fibonacci word

Construction of the string attractors:

$$\mu^\omega(0) = 0100\underline{1}01\underline{0}01001\cdots$$

We have the string attractors $\{1,2\}$, $\{2,3\}$, $\{3,5\}$, $\{5,8\}$, ...

$\longrightarrow$ Fibonacci sequence!

We also have

$$|\mu^0(0)| = 1 \quad |\mu^1(0)| = 2 \quad |\mu^2(0)| = 3 \quad |\mu^3(0)| = 5 \quad \cdots$$

Generalized morphisms

For $k \geq 2$ and for $c_1, \ldots, c_k \geq 0$ with $c_1, c_k \geq 1$, we define

$$\mu_c : \begin{cases} 0 \mapsto 0^{c_1}1 \\ 1 \mapsto 0^{c_2}2 \\ \ldots \\ k-2 \mapsto 0^{c_{k-1}}(k-1) \\ k-1 \mapsto 0^{c_k} \end{cases} \qquad \text{and} \quad \mathbf{x}_c = \mu_c^\omega(0)$$

If $k = 2$ and $c_1 = 1 = c_2$, we obtain the Fibonacci word.

Generalized morphisms

For $k \geq 2$ and for $c_1, \ldots, c_k \geq 0$ with $c_1, c_k \geq 1$, we define

$$\mu_c : \begin{cases} 0 \mapsto 0^{c_1}1 \\ 1 \mapsto 0^{c_2}2 \\ \ldots \\ k-2 \mapsto 0^{c_{k-1}}(k-1) \\ k-1 \mapsto 0^{c_k} \end{cases} \qquad \text{and} \quad \mathbf{x}_c = \mu_c^{\omega}(0)$$

If $k = 2$ and $c_1 = 1 = c_2$, we obtain the Fibonacci word.

Question

Can we use similar techniques to prove that every prefix of $\mathbf{x}_c$ has a "small" string attractor made of lengths of $\mu_c^n(0)$, $n \geq 0$?

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

Examples

$$\mu_{112}: \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$$

$$010201000102010101 \cdots$$

$$\mu_{102}: \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$

$\underline{01}0\underline{2}01000102010101 \cdots$

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$$

$\underline{01}0\underline{2}010$00102010101 $\cdots$

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$$

$\underline{01}0\underline{2}01\underline{00}0102010101 \cdots$

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$$

$$0\underline{1}0\underline{2}010\underline{0}0102010101\cdots$$

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$$

$0\underline{1}0\underline{2}010\underline{0}0102010101 \cdots$

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$$

$$0\underline{1}0\underline{2}010\underline{0}01020\,10\underline{1}\,01 \cdots$$

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$

$0\underline{1}0\underline{2}010\underline{0}0102010\underline{1}01\cdots$

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$$

$$010\underline{2}010\underline{0}010201 0\underline{1}0 1 \cdots$$

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$

$010\underline{2}010\underline{0}0102010\underline{1}01\cdots$

Seems ✓

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$$

$$010\underline{2}010\underline{0}0102010\underline{1}01 \cdots$$

Seems ✓

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 3, 5, 9, 15, \ldots$$

$$012000101012012 \cdots$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$$

$$010\underline{2}010\underline{0}0102010\underline{1}01 \cdots$$

Seems ✓

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 3, 5, 9, 15, \ldots$$

$$\underline{012}000101012012 \cdots$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$

$010\underline{2}010\underline{0}0102010\underline{1}01 \cdots$

Seems ✓

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

$|\mu^n(0)| : 1, 2, 3, 5, 9, 15, \ldots$

$\underline{012}000101012012 \cdots$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$$

$$010\underline{2}010\underline{0}0102010\underline{1}01 \cdots$$

Seems ✓

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 3, 5, 9, 15, \ldots$$

$$\underline{012}\,\underline{00}\,0101012012 \cdots$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$$

$$010\underline{2}010\underline{0}0102010\underline{1}01\cdots$$

Seems ✓

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 3, 5, 9, 15, \ldots$$

$$0\underline{12}0\underline{0}0101012012\cdots$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$$

$$010\underline{2}010\underline{0}0102010\underline{1}01\cdots$$

Seems ✓

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 3, 5, 9, 15, \ldots$$

$$0\underline{12}0\underline{0}0101012012\cdots$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$$

$$010\underline{2}010\underline{0}0102010\underline{1}01 \cdots$$

Seems ✓

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

$$|\mu^n(0)| : 1, 2, 3, 5, 9, 15, \ldots$$

$$0\underline{12}0\underline{0}0101012012 \cdots$$

Examples

$$\mu_{112} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 00 \end{cases}$$

$|\mu^n(0)| : 1, 2, 4, 8, 16, \ldots$

$010\underline{2}010\underline{0}0102010\underline{1}01 \cdots$

Seems ✓

$$\mu_{102} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 2 \\ 2 \mapsto 00 \end{cases}$$

$|\mu^n(0)| : 1, 2, 3, 5, 9, 15, \ldots$

$0\underline{12}0\underline{0}0101012012 \cdots$

✗ for the prefix of length 8

A linear recurrence

We denote U_n the length of $\mu_c^n(0)$.

A linear recurrence

We denote U_n the length of $\mu_c^n(0)$.

Proposition

For all $n \geq 0$, we have

$$\mu_c^n(0) = \begin{cases} \mu_c^{n-1}(0^{c_1}) \cdot \mu_c^{n-2}(0^{c_2}) \cdots \mu_c^{0}(0^{c_n}) \cdot n & \text{if } n \leq k-1; \\ \mu_c^{n-1}(0^{c_1}) \cdot \mu_c^{n-2}(0^{c_2}) \cdots \mu_c^{n-k}(0^{c_k}) & \text{if } n \geq k. \end{cases}$$

In particular,

$$U_n = \begin{cases} c_1 U_{n-1} + c_2 U_{n-2} + \cdots + c_n U_0 + 1 & \text{if } n \leq k-1; \\ c_1 U_{n-1} + c_2 U_{n-2} + \cdots + c_k U_{n-k} & \text{if } n \geq k. \end{cases}$$

A linear recurrence

We denote U_n the length of $\mu_c^n(0)$.

Proposition

For all $n \geq 0$, we have

$$\mu_c^n(0) = \begin{cases} \mu_c^{n-1}(0^{c_1}) \cdot \mu_c^{n-2}(0^{c_2}) \cdots \mu_c^{0}(0^{c_n}) \cdot n & \text{if } n \leq k-1; \\ \mu_c^{n-1}(0^{c_1}) \cdot \mu_c^{n-2}(0^{c_2}) \cdots \mu_c^{n-k}(0^{c_k}) & \text{if } n \geq k. \end{cases}$$

In particular,

$$U_n = \begin{cases} c_1 U_{n-1} + c_2 U_{n-2} + \cdots + c_n U_0 + 1 & \text{if } n \leq k-1; \\ c_1 U_{n-1} + c_2 U_{n-2} + \cdots + c_k U_{n-k} & \text{if } n \geq k. \end{cases}$$

We denote $\mathcal{S}_G$ the greedy numeration system associated with the sequence $(U_n)_{n\geq 0}$.

Another numeration system

We denote $\mathcal{S}_{DM}$ the Dumont-Thomas numeration system associated with $\mathbf{x}_c$, i.e. $\text{rep}_{\mathcal{S}_{DM}}(m)$ is the coefficients of the greedy factorization of $\mathbf{x}_c[1, m]$ into elements of $(\mu_c^n(0))_{n\geq 0}$.

Another numeration system

We denote $\mathcal{S}_{DM}$ the Dumont-Thomas numeration system associated with $\mathbf{x}_c$, i.e. $\text{rep}_{\mathcal{S}_{DM}}(m)$ is the coefficients of the greedy factorization of $\mathbf{x}_c[1, m]$ into elements of $(\mu_c^n(0))_{n \geq 0}$.

Example: For $c = 102$, we have

$$\mathbf{x}_c[1, 8] = 01200010$$

n	$\mu_c^n(0)$
0	0
1	01
2	012
3	01200
4	012000101

so $\text{rep}_{\mathcal{S}_{DM}}(8) =$

Another numeration system

We denote $\mathcal{S}_{DM}$ the Dumont-Thomas numeration system associated with $\mathbf{x}_c$, i.e. $\mathrm{rep}_{\mathcal{S}_{DM}}(m)$ is the coefficients of the greedy factorization of $\mathbf{x}_c[1, m]$ into elements of $(\mu_c^n(0))_{n\geq 0}$.

Example: For $c = 102$, we have

n	$\mu_c^n(0)$
0	0
1	01
2	012
3	01200
4	012000101

$$\mathbf{x}_c[1, 8] = 01200 \cdot 010$$

so $\mathrm{rep}_{\mathcal{S}_{DM}}(8) = \underset{3}{1}$

Another numeration system

We denote $\mathcal{S}_{DM}$ the Dumont-Thomas numeration system associated with $\mathbf{x}_c$, i.e. $\text{rep}_{\mathcal{S}_{DM}}(m)$ is the coefficients of the greedy factorization of $\mathbf{x}_c[1, m]$ into elements of $(\mu_c^n(0))_{n\geq 0}$.

Example: For $c = 102$, we have

$$\mathbf{x}_c[1, 8] = 01200 \cdot 010$$

n	$\mu_c^n(0)$
0	0
1	01
2	012
3	01200
4	012000101

so $\text{rep}_{\mathcal{S}_{DM}}(8) = \underset{3}{1}\underset{2}{0}$

Another numeration system

We denote $\mathcal{S}_{DM}$ the Dumont-Thomas numeration system associated with $\mathbf{x}_c$, i.e. $\text{rep}_{\mathcal{S}_{DM}}(m)$ is the coefficients of the greedy factorization of $\mathbf{x}_c[1, m]$ into elements of $(\mu_c^n(0))_{n \geq 0}$.

Example: For $c = 102$, we have

$$\mathbf{x}_c[1, 8] = 01200 \cdot 01 \cdot 0$$

n	$\mu_c^n(0)$
0	0
1	01
2	012
3	01200
4	012000101

so $\text{rep}_{\mathcal{S}_{DM}}(8) = \underset{3}{1}\underset{2}{0}\underset{1}{1}$

Another numeration system

We denote $\mathcal{S}_{DM}$ the Dumont-Thomas numeration system associated with $\mathbf{x}_c$, i.e. $\text{rep}_{\mathcal{S}_{DM}}(m)$ is the coefficients of the greedy factorization of $\mathbf{x}_c[1, m]$ into elements of $(\mu_c^n(0))_{n\geq 0}$.

Example: For $c = 102$, we have

$$\mathbf{x}_c[1, 8] = 01200 \cdot 01 \cdot 0$$

n	$\mu_c^n(0)$
0	0
1	01
2	012
3	01200
4	012000101

so $\text{rep}_{\mathcal{S}_{DM}}(8) = \underset{3}{1}\underset{2}{0}\underset{1}{1}\underset{0}{1}$.

Another numeration system

We denote $\mathcal{S}_{DM}$ the Dumont-Thomas numeration system associated with $\mathbf{x}_c$, i.e. $\text{rep}_{\mathcal{S}_{DM}}(m)$ is the coefficients of the greedy factorization of $\mathbf{x}_c[1,m]$ into elements of $(\mu_c^n(0))_{n\geq 0}$.

Example: For $c = 102$, we have

$$\mathbf{x}_c[1,8] = 01200 \cdot 01 \cdot 0$$

n	$\mu_c^n(0)$	U_n
0	0	1
1	01	2
2	012	3
3	01200	5
4	012000101	9

so $\text{rep}_{\mathcal{S}_{DM}}(8) = \underset{3\,2\,1\,0}{1011}$. But $\text{rep}_{\mathcal{S}_G}(8) = 1100$!

$\mathcal{S}_G = \mathcal{S}_{DM}$?

Theorem

Let $c = c_1 \cdots c_k$ such that $c_1, c_k \geq 1$. The following are equivalent

- $\mathcal{S}_G = \mathcal{S}_{DM}$;

$\mathcal{S}_G = \mathcal{S}_{DM}$?

Theorem

Let $c = c_1 \cdots c_k$ such that $c_1, c_k \geq 1$. The following are equivalent

- $\mathcal{S}_G = \mathcal{S}_{DM}$;
- *the language of $\mathcal{S}_{DM}$ is given by*

$$\{v_1 \cdots v_n \mid v_1 \neq 0, v_i \cdots v_n \leq_{lex} d_c^\star\}$$

where $d_c^\star = c_1 \cdots c_{k-1}(c_k - 1)c_1 \cdots c_{k-1}(c_k - 1) \cdots$;

$\mathcal{S}_G = \mathcal{S}_{DM}$?

Theorem

Let $c = c_1 \cdots c_k$ such that $c_1, c_k \geq 1$. The following are equivalent

- $\mathcal{S}_G = \mathcal{S}_{DM}$;
- *the language of $\mathcal{S}_{DM}$ is given by*

$$\{v_1 \cdots v_n \mid v_1 \neq 0, v_i \cdots v_n \leq_{lex} d_c^\star\}$$

  where $d_c^\star = c_1 \cdots c_{k-1}(c_k - 1)c_1 \cdots c_{k-1}(c_k - 1) \cdots$;
- *there exists a Parry simple number β such that*

$$c_1 \cdots c_k = (d_\beta(1) - 1)^\ell d_\beta(1);$$

$\mathcal{S}_G = \mathcal{S}_{DM}$?

Theorem

Let $c = c_1 \cdots c_k$ such that $c_1, c_k \geq 1$. The following are equivalent

- $\mathcal{S}_G = \mathcal{S}_{DM}$;
- *the language of $\mathcal{S}_{DM}$ is given by*

$$\{v_1 \cdots v_n \mid v_1 \neq 0, v_i \cdots v_n \leq_{lex} d_c^\star\}$$

where $d_c^\star = c_1 \cdots c_{k-1}(c_k - 1)c_1 \cdots c_{k-1}(c_k - 1) \cdots$;

- *there exists a Parry simple number β such that*

$$c_1 \cdots c_k = (d_\beta(1) - 1)^\ell d_\beta(1);$$

- *for all $i \leq k$,*

$$c_i \cdots c_{k-1}(c_k - 1)c_1 \cdots c_{i-1} \leq c_1 \cdots c_{k-1}(c_k - 1).$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period U_n;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of $\mathbf{x}_c$.*

For $c = 112$, we have

$$010201000102010101\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period* U_n*;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of* $\mathbf{x}_c$*.*

For $c = 112$, we have

$$\underline{01}0\underline{2}01000102010101\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period* U_n*;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of* $\mathbf{x}_c$*.*

For $c = 112$, we have

$$\underline{01}0\underline{2}0100010201010 1\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period U_n;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of $\mathbf{x}_c$.*

For $c = 112$, we have

$$\underline{01}0\underline{2}01000102010101\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period* U_n*;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of* $\mathbf{x}_c$*.*

For $c = 112$, we have

$$\underline{01}0\underline{2}010\color{gray}{00102010101\cdots}$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period U_n;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of $\mathbf{x}_c$.*

For $c = 112$, we have

$$\underline{01}0\underline{2}01\underline{00}0102010101\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period U_n;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of $\mathbf{x}_c$.*

For $c = 112$, we have

$$0\underline{1}0\underline{2}010\underline{0}0102010101\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period U_n;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of $\mathbf{x}_c$.*

For $c = 112$, we have

$$0\underline{1}0\underline{2}010\underline{0}0102010101\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period U_n;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of $\mathbf{x}_c$.*

For $c = 112$, we have

$$0\underline{1}0\underline{2}010\underline{0}01\,02010101 \cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period* U_n;
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of* $\mathbf{x}_c$.

For $c = 112$, we have

$$0\underline{1}0\underline{2}010\underline{0}010\,2010101\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period U_n;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of $\mathbf{x}_c$.*

For $c = 112$, we have

$$0\underline{1}0\underline{2}010\underline{0}0102010101\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period U_n;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of $\mathbf{x}_c$.*

For $c = 112$, we have

$$0\underline{1}0\underline{2}010\underline{0}01020\,10101\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period U_n;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of $\mathbf{x}_c$.*

For $c = 112$, we have

$$0\underline{1}0\underline{2}010\underline{0}01020101 01\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period U_n;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of $\mathbf{x}_c$.*

For $c = 112$, we have

$$0\underline{1}0\underline{2}010\underline{0}0102010101 \cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period* U_n*;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of* $\mathbf{x}_c$.

For $c = 112$, we have

$$0\underline{1}0\underline{2}010\underline{0}01020{\color{blue}10\underline{1}}{\color{gray}01}\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period U_n;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of $\mathbf{x}_c$.*

For $c = 112$, we have

$$0\underline{1}0\underline{2}010\underline{0}0102010\underline{1}01\cdots$$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period* U_n;
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of* $\mathbf{x}_c$.

For $c = 112$, we have

$$\mathbf{0}\underline{\mathbf{1}}0\underline{2}010\underline{0}010201\mathbf{0}\underline{\mathbf{1}}01\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period U_n;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of $\mathbf{x}_c$.*

For $c = 112$, we have

$$0\underline{1}\mathbf{0}\underline{2}010\underline{0}010201 0\underline{1}\mathbf{0}1 \cdots$$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period U_n;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of $\mathbf{x}_c$.*

For $c = 112$, we have

$$0\underline{1}\mathbf{0}\underline{2}010\underline{0}0102010\underline{1}\mathbf{0}1\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period* U_n;
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of* $\mathbf{x}_c$.

For $c = 112$, we have

$$01\underline{1}\mathbf{0}\underline{2}0100\underline{0}010201\,01\underline{1}\mathbf{0}1\cdots$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period* U_n;
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of* $\mathbf{x}_c$.

For $c = 112$, we have

$$\color{red}{010}\underline{2}010\underline{0}010201\color{red}{0\underline{1}0}\color{gray}{1\cdots}$$

✓ when $\mathcal{S}_G = \mathcal{S}_{DM}$

Proposition

If $\mathcal{S}_G = \mathcal{S}_{DM}$, then for all $n \geq 0$

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is periodic of period* U_n*;*
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *is a prefix of* $\mathbf{x}_c$*.*

For $c = 112$, we have

$$010\underline{2}010\underline{0}0102010\underline{1}01\cdots$$

Theorem

If $c_1, \ldots, c_k$ are such that $\mathcal{S}_G = \mathcal{S}_{DM}$, then every prefix of $\mathbf{x}_c$ admits a string attractor composed of k or $k+1$ consecutive elements of $(U_n)_{n\geq 0}$.

Proposition

If $\mathcal{S}_B \neq \mathcal{S}_{DM}$, then there exists n such that

- $\mathbf{x}_c[1, U_{n+1} - 1]$ *is not periodic of period* U_n;
- $\mathbf{x}_c[U_n + 1, U_{n+1} - 1]$ *does not occur before in* $\mathbf{x}_c$.

For $c = 102$, we have

$$0\underline{12}0\underline{0}010101\cdots$$

Open questions

- Given a sequence $(U_n)_{n\geq 0}$, can we find an infinite word $\mathbf{x}$ whose prefixes have string attractors made of U_n's?
- Given a infinite word $\mathbf{x}$, can we find a numeration system $\mathcal{S}$ such that $\mathbf{x}$ is $\mathcal{S}$-automatic and string attractors of $\mathbf{x}$'s prefixes can be easily described using $\mathcal{S}$?

Open questions

- Given a sequence $(U_n)_{n \geq 0}$, can we find an infinite word $\mathbf{x}$ whose prefixes have string attractors made of U_n's?
- Given a infinite word $\mathbf{x}$, can we find a numeration system $\mathcal{S}$ such that $\mathbf{x}$ is $\mathcal{S}$-automatic and string attractors of $\mathbf{x}$'s prefixes can be easily described using $\mathcal{S}$?

Thank you for your attention!