Sturmian words: equivalent definitions

Anna FRID

Aix-Marseille Université, September 2020

Definition

The *(factor) complexity* $p_{\mathbf{u}}(n)$ of an infinite word $\mathbf{u}$ is the number of its distinct factors of length n.

Definition

The *(factor) complexity* $p_{\mathbf{u}}(n)$ of an infinite word $\mathbf{u}$ is the number of its distinct factors of length n.

Theorem (Morse and Hedlund, 1938)

An infinite word $\mathbf{u}$ *either is ultimately periodic, and then its complexity is ultimately constant, or satisfies* $p_{\mathbf{u}}(n) \geq n+1$.

Definition

The *(factor) complexity* $p_{\mathbf{u}}(n)$ of an infinite word $\mathbf{u}$ is the number of its distinct factors of length n.

Theorem (Morse and Hedlund, 1938)

An infinite word $\mathbf{u}$ *either is ultimately periodic, and then its complexity is ultimately constant, or satisfies* $p_{\mathbf{u}}(n) \geq n+1$.

A word $\mathbf{u}$ of complexity $p_{\mathbf{u}}(n) = n+1$ is called *Sturmian*.

Fibonacci word

Example (Fibonacci morphism)

$$\varphi(0) = 01, \varphi(1) = 0$$

$$0 \rightarrow 01 \rightarrow 01\ 0 \rightarrow 010\ 01 \rightarrow 01001\ 010 \rightarrow 01001010\ 01001 \rightarrow \cdots$$

Its fixed point is the *Fibonacci word*

$$\varphi^{\omega}(0) = 0100101001001010010100100101001001 \cdots$$

Fibonacci word

Example (Fibonacci morphism)

$$\varphi(0) = 01, \varphi(1) = 0$$

$$0 \to 01 \to 01\ 0 \to 010\ 01 \to 01001\ 010 \to 01001010\ 01001 \to \cdots$$

Its fixed point is the *Fibonacci word*

$$\varphi^\omega(0) = 0100101001001010010100100101001001\cdots$$

Lemma

The Fibonacci word is Sturmian.

Balanced words

Let $|u|_a$ denote the number of occurrences of a to u

$$|01001|_1 = 2$$

Balanced words

Let $|u|_a$ denote the number of occurrences of a to u

$$|01001|_1 = 2$$

An infinite word $\mathbf{w}$ over $\{0, 1\}$ is said to be *balanced* if for every two its factors x and y of the same length we have

$$\delta(x, y) = ||x|_1 - |y|_1| \leq 1.$$

Balanced words

Let $|u|_a$ denote the number of occurrences of a to u

$$|01001|_1 = 2$$

An infinite word $\mathbf{w}$ over $\{0, 1\}$ is said to be *balanced* if for every two its factors x and y of the same length we have

$$\delta(x, y) = ||x|_1 - |y|_1| \leq 1.$$

Example

$$01001\ 01001\ 01001\ 01001 \cdots$$

Equivalence I

Theorem

A right infinite word is Sturmian if and only if it is aperiodic and balanced.

Equivalence I

Theorem

A right infinite word is Sturmian if and only if it is aperiodic and balanced.

IDEA OF THE PROOF:

- In a balanced set of factors of length n, there are at most $n+1$ elements;

Equivalence I

Theorem

A right infinite word is Sturmian if and only if it is aperiodic and balanced.

IDEA OF THE PROOF:

- In a balanced set of factors of length n, there are at most $n+1$ elements;
- A set of factors F is not balanced $\iff$ there exists a strong bispecial $w|\ 0w0, 1w1 \in F$.

One- and two-sided words

A typical Sturmian word may start for example with

$$001000100100010001001000100100\cdots$$

One- and two-sided words

A typical Sturmian word may start for example with

$$0010001001000100010010001001\cdots$$

Attention:

$$\cdots 00000010000000 \cdots$$

is not considered to be Sturmian, even though its complexity is $n+1$. It is two-sided and "half-periodic".

Slope of a word

Definition

The *slope* of a finite word x over $\{0,1\}$ is

$$\frac{|x|_1}{|x|}.$$

Slope of a word

Definition

The *slope* of a finite word x over $\{0,1\}$ is

$$\frac{|x|_1}{|x|}.$$

The *slope* of an infinite word $\mathbf{w}$ over $\{0,1\}$ is the limit

$$\pi(\mathbf{w}) = \lim_{n\to\infty} \frac{|\mathbf{w}[0..n-1]|_1}{n}.$$

Slope of a word

Definition

The *slope* of a finite word x over $\{0, 1\}$ is

$$\frac{|x|_1}{|x|}.$$

The *slope* of an infinite word $\mathbf{w}$ over $\{0, 1\}$ is the limit

$$\pi(\mathbf{w}) = \lim_{n\to\infty} \frac{|\mathbf{w}[0..n-1]|_1}{n}.$$

Lemma

Every balanced infinite word has a slope.

Slope of a word

Definition

The *slope* of a finite word x over $\{0,1\}$ is

$$\frac{|x|_1}{|x|}.$$

The *slope* of an infinite word $\mathbf{w}$ over $\{0,1\}$ is the limit

$$\pi(\mathbf{w}) = \lim_{n\to\infty} \frac{|\mathbf{w}[0..n-1]|_1}{n}.$$

Lemma

Every balanced infinite word has a slope.

Lemma

A balanced word is periodic if and only if its slope is rational.

Examples

Example

The slope of

$$01001\ 01001\ 01001\ 01001 \cdots$$

is $2/5$.

Examples

Example

The slope of

$$01001\ 01001\ 01001\ 01001 \cdots$$

is $2/5$.

Example

The slope of the Fibonacci word $0\ 1\ 0\ 01\ 010\ 01001\ 01001010 \cdots$ is

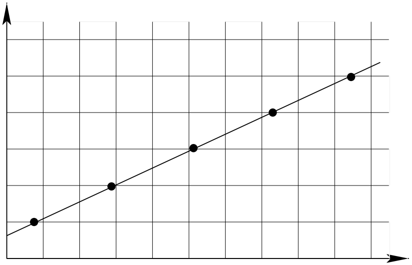
$$\lim_{n\to\infty} \frac{|\varphi^n(0)|_1}{|\varphi^n(0)|} = \lim_{n\to\infty} \frac{F_{n-2}}{F_n} = \frac{1}{\tau^2},$$

where $\tau = (1+\sqrt{5})/2$.

$$\frac{1}{\tau^2} = 0,38 \cdots .$$

Mechanical words

$$y = \sigma x + \rho, \quad 0 \leq \sigma, \rho < 1.$$

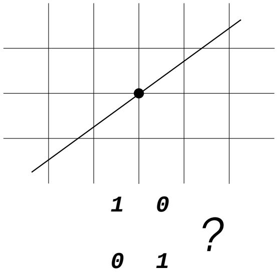


1 0 1 0 0 1 0 1 0 1

$$\mathbf{w} = w[0]w[1]\cdots$$

$$w[n] = \lfloor (n+1)\sigma + \rho \rfloor - \lfloor n\sigma + \rho \rfloor.$$

Important choice



$$w[n] = \lfloor (n+1)\sigma + \rho \rfloor - \lfloor n\sigma + \rho \rfloor.$$

or

$$w[n] = \lceil (n+1)\sigma + \rho \rceil - \lceil n\sigma + \rho \rceil.$$

Mechanical words

Definition

An infinite word $\mathbf{w} = w[0]w[1]\cdots$ over $\{0,1\}$ is *mechanical*, if for all $n \geq 0$ we have

$$w[n] = \lfloor (n+1)\sigma + \rho \rfloor - \lfloor n\sigma + \rho \rfloor,$$

or

$$w[n] = \lceil (n+1)\sigma + \rho \rceil - \lceil n\sigma + \rho \rceil.$$

Mechanical words

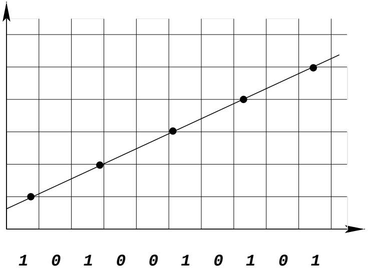
Definition

An infinite word $\mathbf{w} = w[0]w[1]\cdots$ over $\{0,1\}$ is *mechanical*, if for all $n \geq 0$ we have

$$w[n] = \lfloor (n+1)\sigma + \rho \rfloor - \lfloor n\sigma + \rho \rfloor,$$

or

$$w[n] = \lceil (n+1)\sigma + \rho \rceil - \lceil n\sigma + \rho \rceil.$$



Three equivalent definitions

Theorem

For a right infinite word $\mathbf{x}$ *over* $\{0,1\}$*, the following conditions are equivalent:*

- $p_{\mathbf{x}}(n) = n+1 \quad \forall n$;
- x *is a non-periodic balanced word;*
- x *is a mechanical word with an irrational slope* σ.

Three equivalent definitions

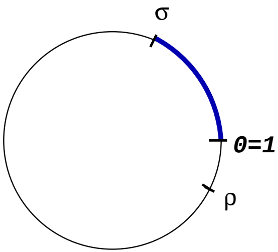
Theorem

For a right infinite word $\mathbf{x}$ *over* $\{0,1\}$, *the following conditions are equivalent:*

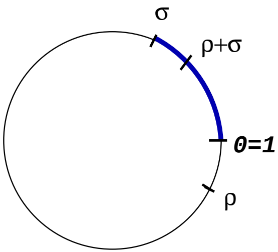
- $p_{\mathbf{x}}(n) = n + 1 \quad \forall n;$
- x *is a non-periodic balanced word;*
- x *is a mechanical word with an irrational slope* σ.

If any of the conditions holds, the word is Sturmian.

Mechanical words and rotations

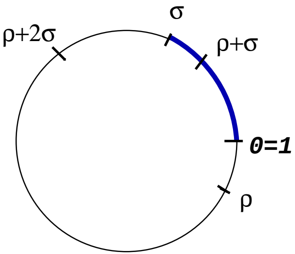


Mechanical words and rotations



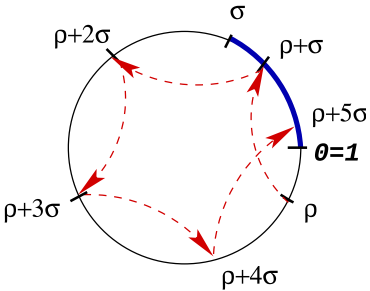
$\mathbf{w} = 1 \cdots$

Mechanical words and rotations



$\mathbf{w} = 10\cdots$

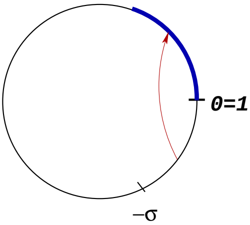
Mechanical words and rotations



$\mathbf{w} = 10001\cdots$

Complexity of rotation words

$w[0] = 1 \iff 1 - \sigma < \rho < 1$

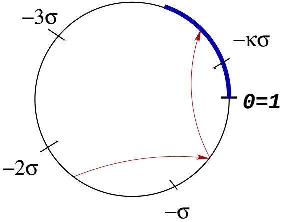


The first symbol is determined by one of two intervals where ρ is located

$p_{\mathbf{w}}(1) = 2$

Complexity of rotation words

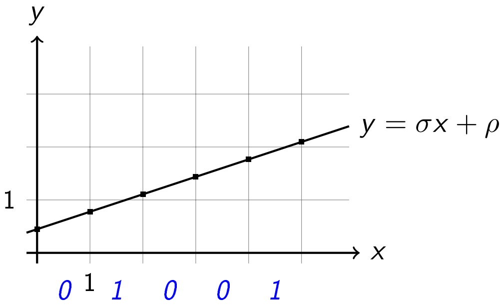
$w[k] = 1 \iff -k\sigma < \rho < -(k+1)\sigma$



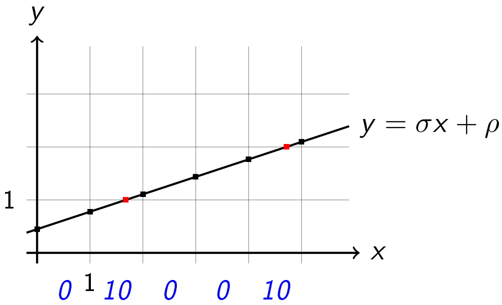
The prefix of length k is determined by one of $k+1$ intervals where ρ is located

$p_w(k) = k + 1.$

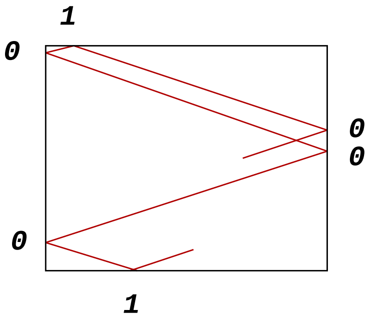
Mechanical vs. billiard definition



Mechanical vs. billiard definition

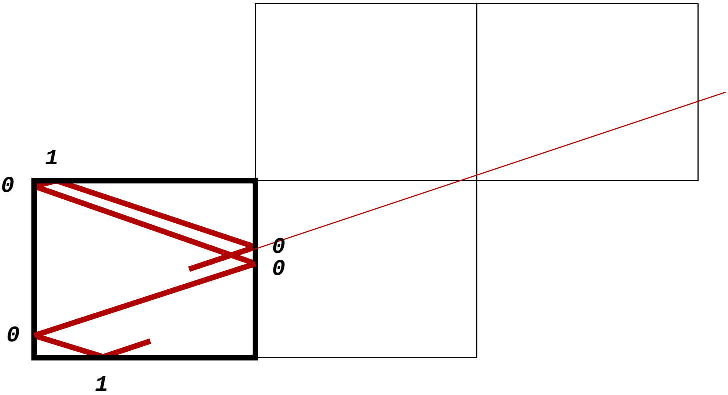


Billiards



$$010001\cdots$$

Billiard words are Sturmian



$$010001\cdots$$

Properties of Sturmian words

Lemma

A Sturmian word is never k-automatic.

Properties of Sturmian words

Lemma

A Sturmian word is never k-automatic.

PROOF. The frequency of 1 in a Sturmian word is irrational (and equal to the slope). In a k-automatic word, this frequency is rational.

Properties of Sturmian words

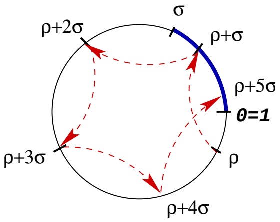
Lemma

The set of factors of a Sturmian word depends only on its slope.

Properties of Sturmian words

Lemma

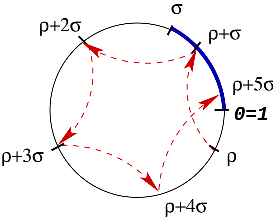
The set of factors of a Sturmian word depends only on its slope.



Properties of Sturmian words

Lemma

The set of factors of a Sturmian word depends only on its slope.



So, for many arguments we may take $\rho = \sigma$. Such Sturmian words are *characteristic*.

Properties of Sturmian words

Lemma

The characteristic word c_σ of slope σ can be constructed from the continued fraction of σ.

Properties of Sturmian words

Lemma

The characteristic word c_σ of slope σ can be constructed from the continued fraction of σ.

Let

$$\sigma = \cfrac{1}{m_1 + 1 + \cfrac{1}{m_2 + \cfrac{1}{m_3 + \cfrac{1}{m_4 + \cdots}}}} = [0, m_1 + 1, m_2, m_3, \cdots].$$

Lemma

The characteristic word c_σ of slope σ can be constructed from the continued fraction of σ.

Let

$$\sigma = \cfrac{1}{m_1 + 1 + \cfrac{1}{m_2 + \cfrac{1}{m_3 + \cfrac{1}{m_4 + \cdots}}}} = [0, m_1 + 1, m_2, m_3, \cdots].$$

Then $c_\sigma = \lim_{n\to\infty} s_n$, where

$$s_{-1} = 1, s_0 = 0, \qquad s_n = s_{n-1}^{m_n} s_{n-2}.$$

Example: the Fibonacci word

The slope of the Fibonacci word is

$$\frac{1}{\tau^2} = [0, 2, 1, 1, 1, 1, \cdots], \text{ where } \tau = \frac{1+\sqrt{5}}{2}.$$

$$\begin{aligned} s_{-1} &= 1 \\ s_0 &= 0 \\ s_1 &= 01 \\ s_2 &= 01\ 0 \\ s_3 &= 010\ 01 \\ s_4 &= 01001\ 010 \end{aligned}$$

Example: the Fibonacci word

The slope of the Fibonacci word is

$$\frac{1}{\tau^2} = [0, 2, 1, 1, 1, 1, \cdots], \text{ where } \tau = \frac{1+\sqrt{5}}{2}.$$

$$\begin{aligned}
s_{-1} &= 1 \\
s_0 &= 0 \\
s_1 &= 01 \\
s_2 &= 01\ 0 \\
s_3 &= 010\ 01 \\
s_4 &= 01001\ 010
\end{aligned}$$

So, it is indeed the Fibonacci word.