S-adic conjecture

November 2, 2010

S-adic sequence

Let a be a letter of a finite alphabet $\mathcal{A}$ and $S = \{\sigma_0, \sigma_1, \ldots, \sigma_{m-1}\}$ be a finite set of morphisms $\sigma_k : \mathcal{A}_k \to \mathcal{A}^*$ with $\mathcal{A}_k \subset \mathcal{A}$.

An infinite word $\mathbf{u}$ over $\mathcal{A}$ is an S-adic sequence if there is a sequence $(\sigma_{i_j})_{j\geq 0}$ of morphisms from S, such that

$$\mathbf{u} = \lim_{n\to\infty} \sigma_{i_0}\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_n}(aaa\cdots) \quad \text{in } \mathcal{A}^{\mathbb{N}}.$$

Any word is an S-adic sequence

Theorem (Cassaigne)

Any infinite word is an S-adic sequence for $\#S = \#\mathcal{A} + 1$.

A Sturmian S-adic sequence

$$\sigma_0 = \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 01 \end{cases} \qquad \sigma_1 = \begin{cases} 0 \mapsto 1 \\ 1 \mapsto 10 \end{cases}$$

A Sturmian S-adic sequence

$$\sigma_0 = \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 01 \end{cases} \qquad \sigma_1 = \begin{cases} 0 \mapsto 1 \\ 1 \mapsto 10 \end{cases}$$

Any sequence

$$\mathbf{u} = \lim_{n\to\infty} \sigma_{i_0}\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_n}(000\cdots).$$

is a Sturmian word.

A Sturmian S-adic sequence

$$\sigma_0 = \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 01 \end{cases} \qquad \sigma_1 = \begin{cases} 0 \mapsto 1 \\ 1 \mapsto 10 \end{cases}$$

Any sequence

$$\mathbf{u} = \lim_{n \to \infty} \sigma_{i_0} \sigma_{i_1} \sigma_{i_2} \cdots \sigma_{i_n}(000 \cdots).$$

is a Sturmian word.
In fact, any Kneading sequence of an irrational α can be written in this form, where $(i_j)_{j \geq 0}$ is determined by the coefficients of the continuous fraction of α.

S-adic Sturmian sequences

Result from: Berthé, Holton, Zamboni, Initial powers of Sturmian sequences. *Acta Arith.* **122**(4):315–347, 2006.

Any Sturmian word is S-adic for $S = \{\tau_0, \tau_0', \tau_1, \tau_1'\}$ where

$$\tau_0 = \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 01 \end{cases} \quad \tau_0' = \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 10 \end{cases} \quad \tau_1 = \begin{cases} 0 \mapsto 10 \\ 1 \mapsto 1 \end{cases} \quad \tau_1' = \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 1. \end{cases}$$

S-adic Sturmian sequences

Result from: Berthé, Holton, Zamboni, Initial powers of Sturmian sequences. *Acta Arith.* **122**(4):315–347, 2006.

Any Sturmian word is S-adic for $S = \{\tau_0, \tau_0', \tau_1, \tau_1'\}$ where

$$\tau_0 = \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 01 \end{cases} \quad \tau_0' = \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 10 \end{cases} \quad \tau_1 = \begin{cases} 0 \mapsto 10 \\ 1 \mapsto 1 \end{cases} \quad \tau_1' = \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 1. \end{cases}$$

If the Sturmian word corresponds to a line $\alpha x + \rho$, the order of the substitutions is given by the coefficients of the continued fraction of α and by the *Ostrowski expansion* of ρ.

Sub-linear complexity

Given $\mathbf{u} = u_0 u_1 u_2 \cdots$, any word $u_i u_{i+1} \cdots u_{i+n-1}$ is a factor of length $n \in \mathbb{N}$.
The (factor) complexity of $\mathbf{u}$ is the function

$$\mathcal{C}_{\mathbf{u}}(n) = \text{ number of factors of } \mathbf{u} \text{ of length } n.$$

Sub-linear complexity

Given $\mathbf{u} = u_0 u_1 u_2 \cdots$, any word $u_i u_{i+1} \cdots u_{i+n-1}$ is a factor of length $n \in \mathbb{N}$.
The (factor) complexity of $\mathbf{u}$ is the function

$$\mathcal{C}_{\mathbf{u}}(n) = \text{ number of factors of } \mathbf{u} \text{ of length } n.$$

An (aperiodic) infinite word $\mathbf{u}$ has a sub-linear complexity if $\mathcal{C}_{\mathbf{u}}(n) \leq an$ for some $a \in \mathbb{R}$.

Sub-linear complexity

Given $\mathbf{u} = u_0 u_1 u_2 \cdots$, any word $u_i u_{i+1} \cdots u_{i+n-1}$ is a factor of length $n \in \mathbb{N}$.
The (factor) complexity of $\mathbf{u}$ is the function

$$\mathcal{C}_{\mathbf{u}}(n) = \text{ number of factors of } \mathbf{u} \text{ of length } n.$$

An (aperiodic) infinite word $\mathbf{u}$ has a sub-linear complexity if $\mathcal{C}_{\mathbf{u}}(n) \leq an$ for some $a \in \mathbb{R}$.

Example

Words with sub-linear complexity: Sturmian words, Arnoux-Rauzy words, fixed point of primitive or uniform substitutions, . . .

S-adic conjecture itself

The S-conjecture is the existence of a (reasonable) condition C such that

> *"**u** has a sub-linear complexity if and only if **u** is S-adic for S satisfying C".*

Cassaigne' result

Result from: J. Cassaigne, Special factors of sequences with linear subword complexity. Developments in language theory, II (Magdeburg, 1995), 25–34, World Sci. Publishing, Singapore, 1996.

Theorem

A word $\mathbf{u}$ *has a sub-linear complexity if and only if the first difference of complexity* $\mathcal{C}_{\mathbf{u}}(n+1) - \mathcal{C}_{\mathbf{u}}(n)$ *is bounded.*

Linearly recurrent words – part 1

A word w is a return word of z in $\mathbf{u}$ if wz is a factor of $\mathbf{u}$, z is a prefix of wz and wz contains exactly two occurrences of z.

Linearly recurrent words – part 1

A word w is a return word of z in $\mathbf{u}$ if wz is a factor of $\mathbf{u}$, z is a prefix of wz and wz contains exactly two occurrences of z.

An infinite word $\mathbf{u}$ is linearly recurrent if any factor z occurs infinitely many times and there is $K \in \mathbb{N}$ such that for any return word w of z we have $|w| \leq K|z|$.

Linearly recurrent words – part 2

Let $\mathbf{u}$ be an S-adic sequence generated by a sequence of morphisms $\sigma_0\sigma_1\sigma_3\cdots$ such that $\sigma_n : \mathcal{A}_{n+1} \to \mathcal{A}_n^*$; $\mathbf{u}$ is called primitive S-adic sequence if there exists $s_0 \in \mathbb{N}$ such that for all r, all $b \in \mathcal{A}_r$ and all $c \in \mathcal{A}_{r+s_0+1}$ the letter b occurs in $\sigma_{r+1}\sigma_{r+2}\cdots\sigma_{r+s_0}(c)$.

Linearly recurrent words – part 2

Let $\mathbf{u}$ be an S-adic sequence generated by a sequence of morphisms $\sigma_0\sigma_1\sigma_3\cdots$ such that $\sigma_n : \mathcal{A}_{n+1} \to \mathcal{A}_n^*$; $\mathbf{u}$ is called primitive S-adic sequence if there exists $s_0 \in \mathbb{N}$ such that for all r, all $b \in \mathcal{A}_r$ and all $c \in \mathcal{A}_{r+s_0+1}$ the letter b occurs in $\sigma_{r+1}\sigma_{r+2}\cdots\sigma_{r+s_0}(c)$.

A morphism $\sigma : \mathcal{A} \to \mathcal{B}^*$ is proper, if there exist two letters $r, l \in \mathcal{B}$ such that $\sigma(a) = lw_ar$, $w_a \in \mathcal{B}^*$, for all $a \in \mathcal{A}$. An S-adic sequence $\mathbf{u}$ is proper, if the morphisms from S are proper.

Linearly recurrent words – part 3

Result from: F. Durand, Corrigendum and addendum to ‘Linearly recurrent subshifts have a finite number of non-periodic factors’. *Ergod. Th. & Dynam. Sys.* (2003), **23**, 663–669.

Theorem

An infinite word $\mathbf{u}$ *is linearly recurrent if and only if it is a primitive and proper* S*-adic sequence.*

Linearly recurrent words – part 3

Result from: F. Durand, Corrigendum and addendum to ‘Linearly recurrent subshifts have a finite number of non-periodic factors’. *Ergod. Th. & Dynam. Sys.* (2003), **23**, 663–669.

Theorem

An infinite word $\mathbf{u}$ *is linearly recurrent if and only if it is a primitive and proper* S*-adic sequence.*

Question

A fixed point of a morphism is linearly recurrent if and only if ***what ?***.

Arnoux-Rauzy words over three letter alphabet

Arnoux-Rauzy words over three letter alphabet $\{0, 1, 2\}$, i.e., with complexity $2n + 1$: all n-segments of the corresponding Rauzy graphs are of the form

$$\sigma_{i_0}\sigma_{i_1} \cdots \sigma_{i_k}(a)$$

for some $k \in \mathbb{N}$ and $a, i_j \in \{0, 1, 2\}$ with

$$\sigma_0 = \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 10 \\ 2 \mapsto 20 \end{cases} \quad \sigma_1 = \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 1 \\ 3 \mapsto 21 \end{cases} \quad \sigma_2 = \begin{cases} 0 \mapsto 02 \\ 1 \mapsto 12 \\ 2 \mapsto 2 \end{cases} .$$

Sufficient condition for sub-linearity – part 1

Proposition

Let $\mathcal{A}$ be a finite alphabet, a be a letter of $\mathcal{A}$, $(\sigma_n : \mathcal{A}_{n+1} \to \mathcal{A}_n^*)_{n\in\mathbb{N}}$ be a sequence of morphisms with $\mathcal{A}_n \subset \mathcal{A}$, $a \in \cap_n \mathcal{A}_n$ and

$$\mathbf{u} = \lim_{n\to\infty} \sigma_0\sigma_1 \cdots \sigma_n(aaa\cdots).$$

Suppose moreover that

$$\lim_{n\to\infty} \inf_{c\in\mathcal{A}_{n+1}} |\sigma_0\sigma_1\cdots\sigma_n(c)| = \infty$$

and there exists a constant D such that

$$\frac{|\sigma_0\sigma_1\cdots\sigma_n\sigma_{n+1}(b)|}{|\sigma_0\sigma_1\cdots\sigma_n(c)|} \leq D$$

for all $b \in \mathcal{A}_{n+2}$, $c \in \mathcal{A}_{n+1}$ and $n \in \mathbb{N}$. Then $\mathcal{C}_{\mathbf{u}}(n) \leq D(\#\mathcal{A})^2 n$.

Sufficient condition for sub-linearity – part 2

Corollary

Let $\mathcal{A}$ be a finite alphabet, a be a letter of $\mathcal{A}$, $(\sigma_n : \mathcal{A} \to \mathcal{A}^*)_{n\in\mathbb{N}}$ be a sequence of k-uniform morphisms with $k > 1$. Then

$$\mathbf{u} = \lim_{n\to\infty} \sigma_0\sigma_1\cdots\sigma_n(aaa\cdots).$$

has a sub-linear complexity $\mathcal{C}_{\mathbf{u}}(n) \leq k(\#\mathcal{A})^2 n$.

(Sort of) necessary condition for sub-linearity

Result from: S. Ferenczi, Rank and symbolic complexity. *Ergod. Th. & Dynam. Sys.* (1996), **16**, 663–682.

Theorem

Let **u** *be a uniformly recurrent word over* $\mathcal{A}$ *with sub-linear complexity. There exists a finite number of morphisms* $\sigma_0, \sigma_1, \ldots, \sigma_{m-1}$ *over* $\mathcal{B} = \{0, 1, \ldots, d-1\}$*, an application* α *from* $\mathcal{B}$ *to* $\mathcal{A}$ *and an infinite sequence* $(i_j)_{j \geq 0}$ *from* $\{0, 1, \ldots, m-1\}^{\mathbb{N}}$ *such that*

$$\lim_{n \to \infty} \inf_{c \in \mathcal{B}} |\sigma_{i_0} \sigma_{i_1} \cdots \sigma_{i_n}(c)| = \infty$$

and any factor of **u** *is a factor of* $\alpha \sigma_{i_0} \sigma_{i_1} \cdots \sigma_{i_n}(0)$ *for some* n*.*

Restriction to fixed points

What is the condition C_1 such that

> *"an infinite (aperiodic) fixed point of a morphism has a sub-linear complexity if and only if the morphism satisfies C_1".*

Sufficient condition for sub-linearity – part 1

Given a morphism σ. The growth function of a letter a is

$$h_a(n) = |\sigma^n(a)|$$

Theorem (Salomaa et al.)

For a non-erasing morphism σ over $\mathcal{A}$ and any $a \in \mathcal{A}$, there exist an integer $e_a \geq 0$ and an algebraic real number ρ_a such that

$$h_a(n) = \Theta(n^{e_a} \rho_a^n) \, .$$

Sufficient condition for sub-linearity – part 1

Given a morphism σ. The growth function of a letter a is

$$h_a(n) = |\sigma^n(a)|$$

Theorem (Salomaa et al.)

For a non-erasing morphism σ over $\mathcal{A}$ and any $a \in \mathcal{A}$, there exist an integer $e_a \geq 0$ and an algebraic real number ρ_a such that

$$h_a(n) = \Theta(n^{e_a} \rho_a^n) .$$

a is bounded if $h_a(n)$ is bounded.

Sufficient condition for sub-linearity – part 2

$$h_a(n) = \Theta(n^{e_a} \rho_a^n)$$

Definition

A morphism σ over $\mathcal{A}$ is said to be

- *non-growing if there is a bounded letter in $\mathcal{A}$*
- *u-exponential if $\rho_a = \rho_b > 1$, $e_a = e_b = 0$ for all $a, b \in \mathcal{A}$,*
- *p-exponential if $\rho_a = \rho_b > 1$ for all $a, b \in \mathcal{A}$ and $e_c > 0$ for some $c \in \mathcal{A}$*
- *e-exponential if $\rho_a > 1$ for all $a \in \mathcal{A}$ and $\rho_b > \rho_c$ for some $b, c \in \mathcal{A}$.*

Sufficient condition for sub-linearity – part 3

Theorem (Ehrenfeucht, Lee, Rozenberg, Pansiot)

Let $\mathbf{u} = \sigma^\omega(a)$ *be an infinite aperiodic word of factor complexity* $\mathcal{C}(n)$.

- *If* σ *is growing, then* $\mathcal{C}(n)$ *is either* $\Theta(n)$, $(n \log\log n)$ *or* $(n \log n)$, *depending on whether* σ *is u-, p- or e-exponential, resp.*
- *If* σ *is not-growing, then either*
  - a) $\mathbf{u}$ *has arbitrarily large factors over the set of bounded letters (i.e.,* σ *is pushy) and then* $\mathcal{C}(n) = \Theta(n^2)$ *or*
  - b) $\mathbf{u}$ *has finitely many factors over the set of bounded letters and then* $\mathcal{C}(n)$ *can be any of* $\Theta(n), \Theta(n \log\log n)$ *or* $\Theta(n \log n)$.

Necessary condition – wanted

It seems that all examples of fixed points with the non-sub-linear complexity contains unbounded powers of words, i.e., the critical exponent is infinite. Such words are called repetitive.

Theorem (Ehrenfeucht, Rozenberg (1983))

It is decidable whether a fixed point of a morphism is repetitive. Any repetitive fixed point $\mathbf{u}$ *is strongly repetitive, i.e., there is a word* w *such that* w^n *is a factor of* $\mathbf{u}$ *for any* n*.*