k-overlap-free binary words

Patrice Séébold

Université Paul Valéry & LIRMM
Montpellier
France

Let $\mathcal{A} = \{a, b\}$.

An *overlap* is a word of the form $xyxyx$ where $x \in \mathcal{A}^+$ and $y \in \mathcal{A}^*$.

A word is *overlap-free* if it does not contain any factor which is an overlap.

Let $\mathcal{A} = \{a, b\}$.

An *overlap* is a word of the form $xyxyx$ where $x \in \mathcal{A}^+$ and $y \in \mathcal{A}^*$.

A word is *overlap-free* if it does not contain any factor which is an overlap.

Let t be the infinite word defined over $\mathcal{A}$ by $t(i) = a$ if $|bin(i)|_1$ is even, $t(i) = b$ otherwise, $i \in \mathbb{N}$.

$$t = abbabaabbaababbabaababbaabbabaabbaaba\ldots$$

Theorem (Thue, 1912)

The word t is an overlap-free bianary infinite word.

A morphism on $\mathcal{A}^*$ is a mapping $f\colon \mathcal{A}^* \to \mathcal{A}^*$ satisfying $f(xy) = f(x)f(y)$ for all $x, y \in \mathcal{A}^*$. Note that f is completely defined by the values $f(a)$ for every letter a on $\mathcal{A}$.
A morphism is called *prolongable on a letter a* if $f(a) = aw$ for some word $w \in \mathcal{A}^+$ such that $f^n(w) \neq \varepsilon$ for all integers $n \geq 1$. By the definition, $f^n(a)$ is a prefix of $f^{n+1}(a)$ for all integers $n \geq 0$ and the sequence $(f^n(a))_{n\geq 0}$ converges to the unique infinite word generated by f from a,

$$f^\omega(a) := \lim_{n\to\infty} f^n(a) = awf(w)f^2(w)\cdots ,$$

which is a fixed point of f.

A morphism $f: \mathcal{A}^* \to \mathcal{A}^*$ generates an infinite word w from a letter $a \in \mathcal{A}$ if there exists $n \in \mathbb{N}$ such that f^n is prolongable on a.

We say that the morphism f *generates an infinite word* if it generates an infinite word from at least one letter.

Morphisms and k-overlap-free binary words

Morphisms

A morphism $f: \mathcal{A}^* \to \mathcal{A}^*$ generates an infinite word w from a letter $a \in \mathcal{A}$ if there exists $n \in \mathbb{N}$ such that f^n is prolongable on a.

We say that the morphism f *generates an infinite word* if it generates an infinite word from at least one letter.

Example

$$\begin{array}{rccc} \mu: & \mathcal{A}^* & \to & \mathcal{A}^* \\ & a & \mapsto & ab \\ & b & \mapsto & ba \end{array}$$

The morphism μ generates two infinite overlap-free words

$$t = \mu^\omega(a) \text{ and } \bar{t} = \mu^\omega(b)$$

Morphisms and k-overlap-free binary words

k-overlap-free binary words

A *k-overlap* is a word of the form $xyxyx$ where x and y are two words with $|x| = k$. A word is *k-overlap-free* if it does not contain k-overlaps.

For example, the word *baabaab* is not overlap-free but it is 2-overlap-free while the word *baabaaba* is not.

Note that a word is 1-overlap-free if and only if it is overlap-free.

A *k-overlap* is a word of the form $xyxyx$ where x and y are two words with $|x| = k$. A word is *k-overlap-free* if it does not contain k-overlaps.

For example, the word *baabaab* is not overlap-free but it is 2-overlap-free while the word *baabaaba* is not.

Note that a word is 1-overlap-free if and only if it is overlap-free.

Theorem

Let $w \in \mathcal{A}^\omega$ be a k-overlap-free binary infinite word for some $k \in \mathbb{N}_+$. The word w is generated by a morphism if and only if $w = t$ or $w = \overline{t}$.

Theorem

Let $w \in \mathcal{A}^\omega$ be a k-overlap-free binary infinite word for some $k \in \mathbb{N}_+$. The word w is generated by a morphism if and only if $w = t$ or $w = \overline{t}$.

Since t and $\overline{t}$ are the only overlap-free binary infinite words that can be generated by morphism, the proof reduces to verify that a binary infinite word which contains an overlap but is k-overlap-free for some integer $k \geq 2$ cannot be generated by a morphism.

From ternary words to binary words

Thue's construction

Let $\mathcal{B} = \{0, 1, 2\}$.

Thue introduced in 1906 the following application

$$\begin{array}{rccc} \delta : & \mathcal{B}^* & \longrightarrow & \mathcal{A}^* \\ & 0 & \longmapsto & a \\ & 1 & \longmapsto & ab \\ & 2 & \longmapsto & abb \end{array}$$

Example

$$t = \delta(m) \text{ where } m = \lambda^\omega(2) \text{ with } \lambda : \begin{array}{ccc} \mathcal{B}^* & \longrightarrow & \mathcal{B}^* \\ 0 & \longmapsto & 1 \\ 1 & \longmapsto & 20 \\ 2 & \longmapsto & 210 \end{array}$$

From ternary words to binary words

Thue's construction

$$t = \delta(m) \text{ where } m = \lambda^\omega(2) \text{ with } \lambda : \begin{array}{ccc} \mathcal{B}^* & \to & \mathcal{B}^* \\ 0 & \mapsto & 1 \\ 1 & \mapsto & 20 \\ 2 & \mapsto & 210 \end{array}$$

Theorem (Thue, 1912)

Let $v \in \mathcal{B}^\omega$. The word $\delta(v)$ is overlap-free if and only if the word v is square-free and does not contain the factor 010 *nor the factor* 212.

From ternary words to binary words

The restricted square property

An infinite word v over $\mathcal{B}$ has the *restricted square property* if for every non-empty factor rr of v

- the word r does not begin nor end with the letter 0,
- in v the factor rr is preceded and followed by the letter 0.

Remark that if a word $v \in \mathcal{B}^\omega$ has the restricted square property then v does not begin with a square, v is overlap-free, and v does not contain 00 as a factor.

Example
The word 21020120 has the restricted square property
The word 21022012120 has the restricted square property
The word 2102201$\underline{11}$210 has not the restricted square property
The word 21$\underline{0202}$ has the restricted square property

From ternary words to binary words

The restricted square property

Let k, p be two integers with $k \geq 2$ and $1 \leq p \leq k-1$.

We associate to (k, p) the following application

$$\begin{array}{rccl} \delta_{k,p}: & \mathcal{B}^* & \longrightarrow & \mathcal{A}^* \\ & 0 & \longmapsto & a^{k-p} \\ & 1 & \longmapsto & a^{k-p}b^p \\ & 2 & \longmapsto & a^{k-p}b^{p+1} \end{array}$$

Remark that $\delta_{2,1} = \delta$

From ternary words to binary words

The restricted square property

Theorem

Let u be an infinite word over $\mathcal{A}$, which does not contain the factor $a^{2(k-p)+1}$, and let v be an infinite word over $\mathcal{B}$, which does not begin with a square, such that $\delta_{k,p}(v) = u$.
The word u is k-overlap-free if and only if v has the restricted square property.

There exist infinite square-free ternary words (for example $m = \lambda^{\omega}(2)$) thus there exist k-overlap-free binary infinite words for every positive integer k.

Partial words are words over $\mathcal{A}$ which, in some positions, contain *holes*, i.e., "do not know"-letters. The holes are represented by $\diamond$. Classical words (called *full* words) are only partial words without holes.

A partial word $u \in \mathcal{A}_\diamond^*$ is a *factor* of a partial word $v \in \mathcal{A}_\diamond^* \cup \mathcal{A}_\diamond^\omega$ if there exist words $x, u' \in \mathcal{A}_\diamond^*$ and $y \in \mathcal{A}_\diamond^* \cup \mathcal{A}_\diamond^\omega$ such that $v = xu'y$ with $u'(i) = u(i)$ whenever neither $u(i)$ nor $u'(i)$ is a hole $\diamond$.

For example, let $u = ab\diamond bba\diamond a$. The length of u is $|u| = 8$, and u contains two holes in positions 3 and 7. Let $v = aa\diamond bb\diamond ba\diamond abbaa\diamond$. The word v contains the word u as a factor in positions 3 and 8. The word u is a suffix of the word v.

k-overlap-free partial words

Words with holes

A partial word is a factor of all the (full) words of the same length in which each $\diamond$ is replaced by any letter of $\mathcal{A}$. We call these (full) words the *completions* of the partial word. In the previous example, if $\mathcal{A} = \{a, b\}$, the partial word u has four completions: *ababbaaa*, *ababbaba*, *abbbbaaa*, and *abbbbaba*.

A partial word is k-overlap-free if all its completions are k-overlap-free.

Example The partial word $abbaba\diamond b$ contains the overlap $baba\diamond$. It is 2-overlap-free because *abbabaab* and *abbababb* are 2-overlap-free.

k-overlap-free partial words

$k = 1$

Theorem (Halava, Harju, Kärki, Séébold, 2008)

An infinite overlap-free binary partial word is either full or of the form $\diamond w$ or $x\diamond w$, where w is an infinite full word and x is a letter. There are infinitely many overlap-free words of each type.

Example Let $t = abbabat'$. The words bat' and bbt' are both overlap-free thus the word $b\diamond t'$ is an infinite overlap-free binary partial word.

k-overlap-free partial words
$k = 1$

Theorem (Halava, Harju, Kärki, Séébold, 2008)

An infinite overlap-free binary partial word is either full or of the form $\diamond w$ or $x\diamond w$, where w is an infinite full word and x is a letter. There are infinitely many overlap-free words of each type.

Example Let $t = abbabat'$. The words bat' and bbt' are both overlap-free thus the word $b\diamond t'$ is an infinite overlap-free binary partial word.

There are no infinite overlap-free binary partial word with at least two holes.

k-overlap-free partial words

$k \geq 2$: Thue again

Consider the following morphism

$$\begin{array}{rccl} \tau : & \mathcal{B}^* & \longrightarrow & \mathcal{B}^* \\ & 0 & \longmapsto & 01201 \\ & 1 & \longmapsto & 020121 \\ & 2 & \longmapsto & 0212021 \end{array}$$

Theorem (Thue, 1912)

The morphism τ preserves square-free words.
The word $\tau^\omega(0)$ is square-free.

k-overlap-free partial words

$k \geq 2$: Thue again

$$\begin{array}{rccl} \tau: & \mathcal{B}^* & \longrightarrow & \mathcal{B}^* \\ & 0 & \longmapsto & 01201 \\ & 1 & \longmapsto & 020121 \\ & 2 & \longmapsto & 0212021 \end{array}$$

The word $\delta(\tau^\omega(0))$ is not overlap-free because $\tau^\omega(0)$ contains the factor 212.
But since $\tau^\omega(0)$ is square-free it has the restricted square property, thus for every $k \geq 2$ the word $\delta_{k,p}(\tau^\omega(0))$ is k-overlap-free.

k-overlap-free partial words

$k \geq 2$

$$\begin{aligned}\tau^{\omega}(0) &= u_1\tau(01)u_2\tau(01)\cdots u_k\tau(01)\cdots, u_i \in \mathcal{B}^+ \\ &= \prod_{k=1}^{\infty} u_k\tau(01) \\ &= \prod_{k=1}^{\infty} u_k 01201\underline{102}0121.\end{aligned}$$

Let $Y = \prod_{k=1}^{\infty} u_k 01202\underline{22}0121$.

Proposition

The word Y has the restricted square property.

k-overlap-free partial words

$k \geq 2$

$$\begin{aligned} \delta_{k,p}(\tau^\omega(0)) &= \prod_{k=1}^{\infty} \delta_{k,p}(u_k0120)\delta_{k,p}(\underline{102})\delta_{k,p}(0121) \\ &= \prod_{k=1}^{\infty} \delta_{k,p}(u_k0120)a^{k-p}b^p a^{k-p} a^{k-p} b^{p+1}\delta_{k,p}(0121) \end{aligned}$$

$$\begin{aligned} \text{and } \delta_{k,p}(Y) &= \prod_{k=1}^{\infty} \delta_{k,p}(u_k0120)\delta_{k,p}(\underline{22})\delta_{k,p}(0121) \\ &= \prod_{k=1}^{\infty} \delta_{k,p}(u_k0120)a^{k-p}b^{p+1} a^{k-p} b^{p+1}\delta_{k,p}(0121) \end{aligned}$$

When $p = k - 1$,
$\delta_{k,p}(\tau^\omega(0)) = \prod_{k=1}^{\infty} \delta_{k,p}(u_k0120)ab^{k-1}\underline{a}ab^k\delta_{k,p}(0121)$ and
$\delta_{k,p}(Y) = \prod_{k=1}^{\infty} \delta_{k,p}(u_k0120)ab^{k-1}\underline{b}ab^k\delta_{k,p}(0121)$.

Theorem

The word $\prod_{k=1}^{\infty} \delta_{k,p}(u_k0120)ab^{k-1}\diamond ab^k\delta_{k,p}(0121)$ is k-overlap-free (but not $(k-1)$-overlap-free).