Bispecial factors in D0L languages

Karel Klouda[1]
karel.klouda@fit.cvut.cz

joint work with Štěpán Starosta[1]

[1]FIT, CTU in Prague

MELA 2013
10th September 2013

Outline

1. What are D0L systems: examples, basic properties.
2. Bispecial factors and why they are interesting.
3. Injective, primitive, pushy and circular D0L systems.
4. Algorithm describing all bispecial factors. It works only for non-repetitive D0L systems!!
5. What about repetitive D0L systems?

Alphabet, (endo)morphism

- An alphabet $\mathcal{A}$ is a finite set of letters,
  - $\{0, 1, 2, 3\}$,
  - $\{a, b, c\}$.

Alphabet, (endo)morphism

- An alphabet $\mathcal{A}$ is a finite set of letters,
  - $\{0, 1, 2, 3\}$,
  - $\{a, b, c\}$.

- A (non-erasing) morphism $\varphi : \mathcal{A}^* \to \mathcal{A}^*$,
  - $0 \to 01, 1 \to 10$,
  - $a \to ab, b \to a$,
  - $0 \to 10, 1 \to 01$.

Alphabet, (endo)morphism

- An alphabet $\mathcal{A}$ is a finite set of letters,
  - $\{0, 1, 2, 3\}$,
  - $\{a, b, c\}$.

- A (non-erasing) morphism $\varphi : \mathcal{A}^* \to \mathcal{A}^*$,
  - $0 \to 01, 1 \to 10$,
  - $a \to ab, b \to a$,
  - $0 \to 10, 1 \to 01$.

- An infinite word $\mathbf{u}$ is a periodic point of φ iff

$$\varphi^{\ell}(\mathbf{u}) = \mathbf{u}$$

for some ℓ. If $\ell = 1$, $\mathbf{u}$ is a fixed point.

D0L system

Definition

A D0L-system is a triplet $G = (\mathcal{A}, \varphi, w)$ where $\mathcal{A}$ is an alphabet, φ a morphism on $\mathcal{A}$, and $w \in \mathcal{A}^+$ is the axiom.
The sequence of G:

$$L(G) = \{w_0 = w, w_1 = \varphi(w_0), w_2 = \varphi(w_1), \ldots\}.$$

All factors of $w_1, w_2, \ldots$ form the language of G, denoted as $sub(L(G))$.

D0L system

Definition

A D0L-system is a triplet $G = (\mathcal{A}, \varphi, w)$ where $\mathcal{A}$ is an alphabet, φ a morphism on $\mathcal{A}$, and $w \in \mathcal{A}^+$ is the axiom.
The sequence of G:

$$L(G) = \{w_0 = w, w_1 = \varphi(w_0), w_2 = \varphi(w_1), \ldots\}.$$

All factors of $w_1, w_2, \ldots$ form the language of G, denoted as $sub(L(G))$.

$G = (\{0,1,2,3,4\}, \varphi, 013)$ with $\varphi = (0310, 212, 121, 4, 3)$:

$$w_0 = 013$$
$$w_1 = 0310\,212\,4$$
$$w_2 = 031041210310\,121212121\,3$$
$$w_3 = 0310412103103212121212\cdots0310\,2121212121\cdots212\,4$$
$$\vdots$$

D0L system: possible behaviour

$\varphi = (0310, 212, 121, 4, 3)$

D0L system: possible behaviour

$\varphi = (0310, 212, 121, 4, 3)$

Finite system: $G = (\mathcal{A}, \varphi, 4)$:

$$L(G) = \{4, 3, 4, 3, \ldots\}, \quad \text{sub}(L(G)) = \{3, 4\}.$$

Definition

Given a morphism φ on $\mathcal{A}$. A letter $a \in \mathcal{A}$ is bounded if the language of $(\mathcal{A}, \varphi, a)$ is finite.

D0L system: possible behaviour

$\varphi = (0310, 212, 121, 4, 3)$

Finite system: $G = (\mathcal{A}, \varphi, 4)$:

$$L(G) = \{4, 3, 4, 3, \ldots\}, \quad \mathrm{sub}(L(G)) = \{3, 4\}.$$

Definition

Given a morphism φ on $\mathcal{A}$. A letter $a \in \mathcal{A}$ is bounded if the language of $(\mathcal{A}, \varphi, a)$ is finite.

Periodic system $G = (\mathcal{A}, \varphi, 2)$.

D0L system: possible behaviour

$\varphi = (0310, 212, 121, 4, 3)$

Finite system: $G = (\mathcal{A}, \varphi, 4)$:

$$L(G) = \{4, 3, 4, 3, \ldots\}, \quad \mathrm{sub}(L(G)) = \{3, 4\}.$$

Definition

Given a morphism φ on $\mathcal{A}$. A letter $a \in \mathcal{A}$ is bounded if the language of $(\mathcal{A}, \varphi, a)$ is finite.

Periodic system $G = (\mathcal{A}, \varphi, 2)$.

Aperiodic system $G = (\mathcal{A}, \varphi, 0)$, in fact, the language of G is the language of the fixed point of φ starting with 0.

Bispecial factors (a.k.a. BS factors)

Definition

For a word v in a language $\mathcal{L} \subset \mathcal{A}^*$ we define the set of left extensions

$$Lext(v) := \{a \in \mathcal{A} \mid av \in \mathcal{L}\}.$$

If $\#Lext(v) > 1$, then v is said to be left special (LS) factor. Right special (RS) factors are defined. If v is both LS and RS, it is called bispecial.

Bispecial factors (a.k.a. BS factors)

Definition

For a word v in a language $\mathcal{L} \subset \mathcal{A}^*$ *we define the set of left extensions*

$$Lext(v) := \{a \in \mathcal{A} \mid av \in \mathcal{L}\}.$$

If $\#Lext(v) > 1$, *then v is said to be left special (LS) factor. Right special (RS) factors are defined. If v is both LS and RS, it is called bispecial.*

Example

Thue-Morse: $\varphi : (01, 10)$, *fixed point* $\varphi^\infty(0) = 0110\underline{100}1\underline{100}1011010010110\cdots$

100 *is LS but not RS*

Bispecial factors (a.k.a. BS factors)

Definition

For a word v in a language $\mathcal{L} \subset \mathcal{A}^*$ *we define the set of left extensions*

$$Lext(v) := \{a \in \mathcal{A} \mid av \in \mathcal{L}\}.$$

If $\#Lext(v) > 1$, *then v is said to be left special (LS) factor. Right special (RS) factors are defined. If v is both LS and RS, it is called bispecial.*

Example

Thue-Morse: $\varphi : (01, 10)$, *fixed point* $\varphi^\infty(0) = 0110\underline{1001}\,\underline{1001}011010010110\cdots$

100 *is LS but not RS*

1001 *is BS*

BS factors: why bother? – Factor complexity

The factor complexity of a language $\mathcal{L}$:

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

BS factors: why bother? – Factor complexity

The factor complexity of a language $\mathcal{L}$:

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

We have

$$\mathcal{C}(n+1) - \mathcal{C}(n) = \sum_{\substack{v \in \mathcal{L}_n \\ v \text{ is LS}}} (\#\mathrm{Lext}(v) - 1).$$

BS factors: why bother? – Factor complexity

The factor complexity of a language $\mathcal{L}$:

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

We have

$$\mathcal{C}(n+1) - \mathcal{C}(n) = \sum_{\substack{v \in \mathcal{L}_n \\ v \text{ is LS}}} (\#\mathrm{Lext}(v) - 1).$$

Complete knowledge of all LS factors along with the number of their left extensions allows us to evaluate $\mathcal{C}_{\mathcal{L}}(n)$.

Factor complexity of D0L languages

Theorem (Pansiot (1984))

Let $\mathcal{L}$ be a language of a D0L system, then one of the following holds:

(i) $\mathcal{C}_{\mathcal{L}}(n) = \Theta(1)$,

(ii) $\mathcal{C}_{\mathcal{L}}(n) = \Theta(n)$,

(iii) $\mathcal{C}_{\mathcal{L}}(n) = \Theta(n \log \log n)$,

(iv) $\mathcal{C}_{\mathcal{L}}(n) = \Theta(n \log n)$,

(v) $\mathcal{C}_{\mathcal{L}}(n) = \Theta(n^2)$.

Factor complexity of D0L languages

Theorem (Pansiot (1984))

Let $\mathcal{L}$ be a language of a D0L system, then one of the following holds:

(i) $\mathcal{C}_{\mathcal{L}}(n) = \Theta(1)$,

(ii) $\mathcal{C}_{\mathcal{L}}(n) = \Theta(n)$,

(iii) $\mathcal{C}_{\mathcal{L}}(n) = \Theta(n \log\log n)$,

(iv) $\mathcal{C}_{\mathcal{L}}(n) = \Theta(n \log n)$,

(v) $\mathcal{C}_{\mathcal{L}}(n) = \Theta(n^2)$.

Theorem (Salomaa, Soittola (1978))

The sequence $|\varphi^n(a)|$ is either bounded or it grows like $n^{x_a} y_a^n$ with $y_a > 1$ and $x_a \in \mathbb{N}$.

BS factors: why bother? – Maximal repetitions

Example

Let $w = 123$ *and* $v = 12312312312 = (123)^3 12$*, then* $r = \frac{|v|}{|w|} = \frac{11}{3} = 3 + \frac{2}{3}$ *and so* v *is* $\frac{11}{3}$*-power of* w*.*

BS factors: why bother? – Maximal repetitions

Example

Let $w = 123$ *and* $v = 12312312312 = (123)^3 12$*, then* $r = \frac{|v|}{|w|} = \frac{11}{3} = 3 + \frac{2}{3}$ *and so* v *is* $\frac{11}{3}$*-power of* w*.*

Definition

The index of w *in a language* $\mathcal{L}$ *is the number*

$$ind(w) = \sup\{r \in \mathbb{Q} \mid w^r \text{ is a factor of } \mathcal{L}\}.$$

The critical exponent of a language $\mathcal{L}$ *is given by*

$$E(\mathcal{L}) = \sup\{ind(w) \mid w \in \mathcal{L}\}.$$

BS factors: why bother? – Maximal repetitions

Example

Let $w = 123$ *and* $v = 12312312312 = (123)^3 12$*, then* $r = \frac{|v|}{|w|} = \frac{11}{3} = 3 + \frac{2}{3}$ *and so* v *is* $\frac{11}{3}$*-power of* w*.*

Definition

The index of w *in a language* $\mathcal{L}$ *is the number*

$$ind(w) = \sup\{r \in \mathbb{Q} \mid w^r \text{ is a factor of } \mathcal{L}\}.$$

The critical exponent of a language $\mathcal{L}$ *is given by*

$$E(\mathcal{L}) = \sup\{ind(w) \mid w \in \mathcal{L}\}.$$

(i) $1 < \mathrm{E}(\mathcal{L}) \leq \infty$,

(ii) $\mathrm{E}(\mathcal{L}) \to 1 \quad \Rightarrow \quad \#\mathcal{A} \to \infty$,

(iii) for all real $x > 1$ there is an $\mathcal{L}$ with $\mathrm{E}(\mathcal{L}) = x$ (Krieger, Shallit (2007)).

BS factors: why bother? – Maximal repetitions

A word $\overline{w}$ is conjugate of a word w if there are words u and v such that $w = uv$ and $\overline{w} = vu$.

BS factors: why bother? – Maximal repetitions

A word $\overline{w}$ is conjugate of a word w if there are words u and v such that $w = uv$ and $\overline{w} = vu$.

If w have the maximal index among its conjugates and

$$w^{\mathrm{ind}(w)} = w^k w' \qquad \text{with } \mathrm{ind}(w) > 1,$$

then all the following factors are bispecial:

$$w', ww', \ldots, w^{k-1}w'.$$

BS factors: why bother? – Maximal repetitions

A word $\overline{w}$ is conjugate of a word w if there are words u and v such that $w = uv$ and $\overline{w} = vu$.

If w have the maximal index among its conjugates and

$$w^{\text{ind}(w)} = w^k w' \qquad \text{with } \text{ind}(w) > 1,$$

then all the following factors are bispecial:

$$w', ww', \dots, w^{k-1}w'.$$

Definition

Denote by $\mathcal{B}(\mathcal{L})$ the set of ordered pairs (v, w) of factors satisfying the following conditions:

(i) *v is a BS factor,*

(ii) *wv is a power of w in $\mathcal{L}$.*

BS factors: why bother? – Maximal repetitions

A word $\overline{w}$ is conjugate of a word w if there are words u and v such that $w = uv$ and $\overline{w} = vu$.

If w have the maximal index among its conjugates and

$$w^{\mathrm{ind}(w)} = w^k w' \qquad \text{with } \mathrm{ind}(w) > 1,$$

then all the following factors are bispecial:

$$w', ww', \ldots, w^{k-1}w'.$$

Definition

Denote by $\mathcal{B}(\mathcal{L})$ the set of ordered pairs (v, w) of factors satisfying the following conditions:

(i) *v is a BS factor,*

(ii) *wv is a power of w in $\mathcal{L}$.*

$$\mathrm{E}(\mathcal{L}) = \sup\{\mathrm{ind}(w) \mid (v, w) \in \mathcal{B}(\mathcal{L}) \text{ for some BS factor } v\}.$$

Critical exponent of fixed points

Theorem (Krieger (2009))

Let φ be a non-erasing morphism defined over a finite alphabet $\mathcal{A}$, $\mathbf{u} = \varphi^\infty(0)$. Let M_φ be the incident matrix of φ and $\lambda, \lambda_1, \ldots, \lambda_\ell$ be its eigenvalues. Suppose $E(\mathbf{u}) < \infty$. Then

$$E(\mathbf{u}) \in \mathbb{Q}[\lambda, \lambda_1, \ldots, \lambda_\ell]$$

is algebraic of degree at most $\#\mathcal{A}$.

Injective morphisms

Definition

A morphism φ on $\mathcal{A}$ is injective if for all $v_1, v_2 \in \mathcal{A}^*$

$$\varphi(v_1) = \varphi(v_2) \quad \Longrightarrow \quad v_1 = v_2\,,$$

i.e. $\{\varphi(a) \mid a \in \mathcal{A}\}$ is a code.

Injective morphisms

Definition

A morphism φ on $\mathcal{A}$ is injective if for all $v_1, v_2 \in \mathcal{A}^*$

$$\varphi(v_1) = \varphi(v_2) \quad \Longrightarrow \quad v_1 = v_2\,,$$

i.e. $\{\varphi(a) \mid a \in \mathcal{A}\}$ is a code.

It is easy to decide whether a given morphism is injective.

Injective morphisms

Definition

A morphism φ on $\mathcal{A}$ is injective on a language $\mathcal{L}$ if for all $v_1, v_2 \in \mathcal{L}$

$$\varphi(v_1) = \varphi(v_2) \quad \Longrightarrow \quad v_1 = v_2 \, .$$

Injective morphisms

Definition

A morphism φ *on* $\mathcal{A}$ *is injective on a language* $\mathcal{L}$ *if for all* $v_1, v_2 \in \mathcal{L}$

$$\varphi(v_1) = \varphi(v_2) \quad \Longrightarrow \quad v_1 = v_2 \,.$$

Example

The morphism $\varphi = (aca, badc, acab, adc)$ *is not injective, as* $\varphi(ab) = aca\,badc = acab\,adc = \varphi(cd)$.

Injective morphisms

Definition

A morphism φ on $\mathcal{A}$ is injective on a language $\mathcal{L}$ if for all $v_1, v_2 \in \mathcal{L}$

$$\varphi(v_1) = \varphi(v_2) \quad \Longrightarrow \quad v_1 = v_2\,.$$

Example

The morphism $\varphi = (aca, badc, acab, adc)$ is not injective, as $\varphi(ab) = aca\,badc = acab\,adc = \varphi(cd)$.
But it is injective on the language of the D0L-system $(\{a, b, c, d\}, \varphi, a)$ as the factor cd is not included.

Injective morphisms

Definition

A morphism φ *on* $\mathcal{A}$ *is injective on a language* $\mathcal{L}$ *if for all* $v_1, v_2 \in \mathcal{L}$

$$\varphi(v_1) = \varphi(v_2) \quad \Longrightarrow \quad v_1 = v_2\,.$$

Example

The morphism $\varphi = (aca, badc, acab, adc)$ *is not injective, as* $\varphi(ab) = aca\,badc = acab\,adc = \varphi(cd)$.
But it is injective on the language of the D0L-system $(\{a, b, c, d\}, \varphi, a)$ *as the factor* cd *is not included.*

Open Problem

Given a D0L system $(\mathcal{A}, \varphi, w)$. *Decide whether* φ *is injective on its language.*

Primitive morphisms

Definition

A morphism φ on $\mathcal{A}$ is primitive if there exists $k \in \mathbb{N}$ such that for any pair of (possibly equal) letters $a, b \in \mathcal{A}$ the word $\varphi^k(a)$ contains b as its factor.

If φ is primitive, the D0L systems $(\mathcal{A}, \varphi, a)$ and $(\mathcal{A}, \varphi, b)$ have the same language for all $a, b \in \mathcal{A}$.

Primitive morphisms

Definition

A morphism φ on $\mathcal{A}$ is primitive if there exists $k \in \mathbb{N}$ such that for any pair of (possibly equal) letters $a, b \in \mathcal{A}$ the word $\varphi^k(a)$ contains b as its factor.

If φ is primitive, the D0L systems $(\mathcal{A}, \varphi, a)$ and $(\mathcal{A}, \varphi, b)$ have the same language for all $a, b \in \mathcal{A}$.

Example

Thue-Morse $(01, 10)$ *is primitive (with* $k = 1$*), Fibonacci* $(01, 0)$ *as well (with* $k = 2$*).*
The morphism $(0310, 212, 121, 4, 3)$ *is not primitive.*

Repetitive D0L system

Definition

A D0L system G is repetitive if for all $k \in \mathbb{N}$, there exists a word v such that v^k is in the language of G.
It is strongly repetitive if there is a word v such that v^k is in the language of G.

Repetitive D0L system

Definition

A DOL system G is repetitive if for all $k \in \mathbb{N}$, there exists a word v such that v^k is in the language of G.
It is strongly repetitive if there is a word v such that v^k is in the language of G.

Example

The D0L system $(\{0, 1, 2, 3, 4\}, \varphi, 0)$ *with* $\varphi = (0310, 212, 121, 4, 3)$ *is strongly repetitive with* $v = 21$.

Repetitive D0L system

Definition

A DOL system G is repetitive if for all $k \in \mathbb{N}$, there exists a word v such that v^k is in the language of G.
It is strongly repetitive if there is a word v such that v^k is in the language of G.

Example

The D0L system $(\{0,1,2,3,4\}, \varphi, 0)$ *with* $\varphi = (0310, 212, 121, 4, 3)$ *is strongly repetitive with* $v = 21$.

Theorem (Ehrenfeucht, Rozenberg (1983))

Every repetitive D0L system is strongly repetitive.

Pushy D0L system

For a given D0L system $G = (\mathcal{A}, \varphi, w)$ denote $\mathcal{A}_0$ the set of all bounded letters.

Definition

A DOL system G is pushy, if its language contains infinite number of factors over $\mathcal{A}_0$.
If is non-pushy, we denote C_P the length of the longest factor over $\mathcal{A}_0$.

Pushy D0L system

For a given D0L system $G = (\mathcal{A}, \varphi, w)$ denote $\mathcal{A}_0$ the set of all bounded letters.

Definition

A DOL system G is pushy, if its language contains infinite number of factors over $\mathcal{A}_0$.
If is non-pushy, we denote C_P the length of the longest factor over $\mathcal{A}_0$.

Example

The D0L system $(\{0, 1, 2, 3, 4\}, \varphi, 0)$ *with* $\varphi = (0310, 212, 121, 4, 3)$. *The bounded letters are* $\mathcal{A}_0 = \{3, 4\}$. *The system is non-pushy and* $C_P = 1$

Pushy D0L system

For a given D0L system $G = (\mathcal{A}, \varphi, w)$ denote $\mathcal{A}_0$ the set of all bounded letters.

Definition

A DOL system G is pushy, if its language contains infinite number of factors over $\mathcal{A}_0$.
If is non-pushy, we denote C_P the length of the longest factor over $\mathcal{A}_0$.

Example

The D0L system $(\{0,1,2,3,4\}, \varphi, 0)$ with $\varphi = (0310, 212, 121, 4, 3)$. The bounded letters are $\mathcal{A}_0 = \{3,4\}$. The system is non-pushy and $C_P = 1$

Example

Consider again the D0L system $(\{0,1,2,3,4\}, \varphi, 0)$ with $\varphi = (03103, 212, 121, 4, 3)$. The bounded letters are $\mathcal{A}_0 = \{3,4\}$. The system is pushy as $(34)^k$ is a factor for all $k \in \mathbb{N}$.

Pushy D0L system: what is known

- It is decidable whether a DOL system is pushy and C_P is effectively computable (Ehrenfeucht, Rozenberg (1983)).
  - Pushy iff edge condition: there exist $a \in \mathcal{A}$, $k \in \mathbb{N}^+$, $v \in \mathcal{A}^*$ and $u \in \mathcal{A}_0^+$ such that $\varphi^k(a) = vau$ or $\varphi^k(a) = uav$.

Pushy D0L system: what is known

- It is decidable whether a DOL system is pushy and C_P is effectively computable (Ehrenfeucht, Rozenberg (1983)).
  - Pushy iff edge condition: there exist $a \in \mathcal{A}$, $k \in \mathbb{N}^+$, $v \in \mathcal{A}^*$ and $u \in \mathcal{A}_0^+$ such that $\varphi^k(a) = vau$ or $\varphi^k(a) = uav$.
- An algorithm based on a simple graphs given by (KK, ŠS (2013)).
  - Graph on unbounded letters: there is a directed edge from a to b with label u if $\varphi(a) = vbu$ with $v \in \mathcal{A}^*$ and $u \in \mathcal{A}_0^*$.
  - Pushy iff there is a cycle with a non empty label.

Pushy D0L system: what is known

- It is decidable whether a DOL system is pushy and C_P is effectively computable (Ehrenfeucht, Rozenberg (1983)).
  - Pushy iff edge condition: there exist $a \in \mathcal{A}$, $k \in \mathbb{N}^+$, $v \in \mathcal{A}^*$ and $u \in \mathcal{A}_0^+$ such that $\varphi^k(a) = vau$ or $\varphi^k(a) = uav$.
- An algorithm based on a simple graphs given by (KK, ŠS (2013)).
  - Graph on unbounded letters: there is a directed edge from a to b with label u if $\varphi(a) = vbu$ with $v \in \mathcal{A}^*$ and $u \in \mathcal{A}_0^*$.
  - Pushy iff there is a cycle with a non empty label.

Theorem (Cassaigne (2010))

If G is a non-erasing pushy D0L system, then there exist $K \in \mathbb{N}$ and a finite set $\mathcal{U}$ of words from $\mathcal{A}_0^+$ such that every factor from $sub(L(G)) \cap \mathcal{A}_0^+$ is of one of the following three forms:

(i) w_1,

(ii) $w_1 u_1^{k_1} w_2$,

(iii) $w_1 u_1^{k_1} w_2 u_2^{k_2} w_3$,

where $u_1, u_2 \in \mathcal{U}$, $|w_j| < K$ for all $j \in \{1, 2, 3\}$, and $k_1, k_2 \in \mathbb{N}^+$.

Circular D0L systems

Definition

Let $G = (\mathcal{A}, \varphi, w)$ be a D0L system with φ injective and let w be a factor of its language $\mathcal{L}$. An ordered pair of factors (w_1, w_2) is called a synchronizing point of w if $w = w_1 w_2$ and

$$\forall v_1, v_2 \in \mathcal{A}^*, (v_1 w v_2 \in \varphi(\mathcal{L}) \Rightarrow v_1 w_1 \in \varphi(\mathcal{L}) \text{ and } v_2 w_2 \in \varphi(\mathcal{L}).$$

A factor with at least one synchronizing point is called non-synchronized.

Circular D0L systems

Definition

Let $G = (\mathcal{A}, \varphi, w)$ be a D0L system with φ injective and let w be a factor of its language $\mathcal{L}$. An ordered pair of factors (w_1, w_2) is called a synchronizing point of w if $w = w_1 w_2$ and

$$\forall v_1, v_2 \in \mathcal{A}^*, (v_1 w v_2 \in \varphi(\mathcal{L}) \Rightarrow v_1 w_1 \in \varphi(\mathcal{L}) \text{ and } v_2 w_2 \in \varphi(\mathcal{L}).$$

A factor with at least one synchronizing point is called non-synchronized.

Definition

A D0L system is circular if there is only a finite number of non-synchronized factors. The smallest C_S such that $|v| > C_S \Longrightarrow v$ is non-synchronized. The number C_S is called synchronizing delay.

Circular D0L systems

Definition

Let $G = (\mathcal{A}, \varphi, w)$ be a D0L system with φ injective and let w be a factor of its language $\mathcal{L}$. An ordered pair of factors (w_1, w_2) is called a synchronizing point of w if $w = w_1 w_2$ and

$$\forall v_1, v_2 \in \mathcal{A}^*, (v_1 w v_2 \in \varphi(\mathcal{L}) \Rightarrow v_1 w_1 \in \varphi(\mathcal{L}) \text{ and } v_2 w_2 \in \varphi(\mathcal{L}).$$

A factor with at least one synchronizing point is called non-synchronized.

Definition

A D0L system is circular if there is only a finite number of non-synchronized factors. The smallest C_S such that $|v| > C_S \implies v$ is non-synchronized. The number C_S is called synchronizing delay.

Open Problem

Given a circular D0L system, is there some bound on the value of the synchronizing delay?

Circular D0L systems: what is (un)known

Theorem (Mignosi, Séébold (1993))

A D0L system with an injective morphism is repetitive iff it is circular and non-pushy.

Circular D0L systems: what is (un)known

Theorem (Mignosi, Séébold (1993))

A D0L system with an injective morphism is repetitive iff it is circular and non-pushy.

Conjecture

If G is a non-circular D0L system, then there exist $K \in \mathbb{N}$ and a finite set $\mathcal{U}$ of words from $\mathcal{A}^+$ such that every non-synchronized factor from $sub(L(G)) \cap \mathcal{A}^+$ is of one of the following three forms:

(i) w_1,

(ii) $w_1 u_1^{k_1} w_2$, *(maybe:* $w_1 u_1^{k_1} w_2 u_2^{k_2} w_3$*),*

where $u_1, u_2 \in \mathcal{U}$, $|w_j| < K$ for all $j \in \{1,2,3\}$, and $k_1, k_2 \in \mathbb{N}^+$.

Circular D0L systems: what is (un)known

Theorem (Mignosi, Séébold (1993))

A D0L system with an injective morphism is repetitive iff it is circular and non-pushy.

Conjecture

If G is a non-circular D0L system, then there exist $K \in \mathbb{N}$ and a finite set $\mathcal{U}$ of words from $\mathcal{A}^+$ such that every non-synchronized factor from $sub(L(G)) \cap \mathcal{A}^+$ is of one of the following three forms:

(i) w_1,

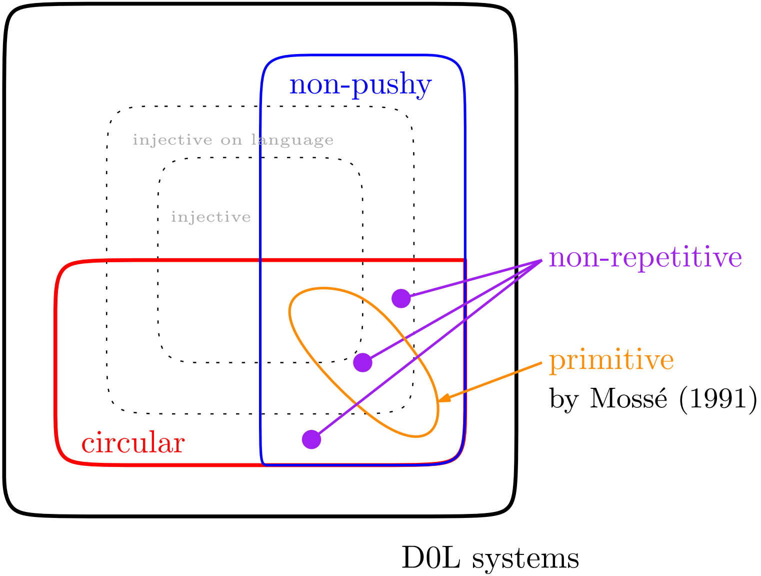
(ii) $w_1 u_1^{k_1} w_2$, *(maybe:* $w_1 u_1^{k_1} w_2 u_2^{k_2} w_3$*)*,

where $u_1, u_2 \in \mathcal{U}$, $|w_j| < K$ *for all* $j \in \{1, 2, 3\}$, *and* $k_1, k_2 \in \mathbb{N}^+$.

Witnesses:

- (Mignosi, Séébold (1993)): if there is infinite non-synchronized factor, then there is an infinite repetition.
- (KK, SŠ (2013)): if a D0L system is not circular, then there are infinite non-synchronized factors of the form (*ii*).

D0L systems



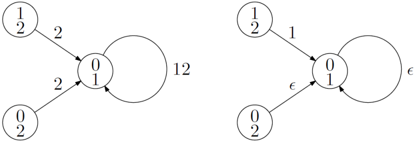
non-pushy
injective on language
injective
non-repetitive
primitive
by Mossé (1991)
circular
D0L systems

Example: intuitive definition of f-image

Definition

In a language $\mathcal{L}$ $((w_1, w_3), v, (w_2, w_4))$ is called a BS triplet if w_i an v are empty words and $w_1 v w_2, w_3 v w_4 \in \mathcal{L}$ or $w_1 v w_4, w_3 v w_2 \in \mathcal{L}$.

Intuitive definition of f-image of BS triplets using graph of prolongations for $\varphi = (012, 112, 102)$:

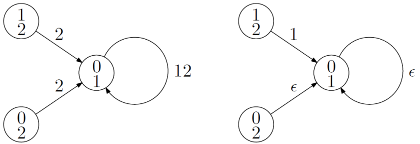


Example: intuitive definition of f-image

Definition

In a language $\mathcal{L}$ $((w_1, w_3), v, (w_2, w_4))$ is called a BS triplet if w_i an v are empty words and $w_1 v w_2, w_3 v w_4 \in \mathcal{L}$ or $w_1 v w_4, w_3 v w_2 \in \mathcal{L}$.

Intuitive definition of f-image of BS triplets using graph of prolongations for $\varphi = (012, 112, 102)$:

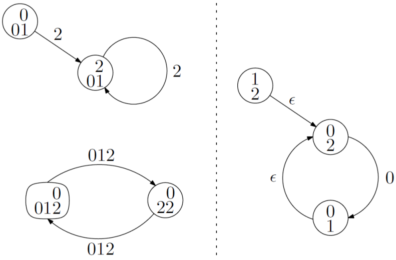


All BS triplets are f^n-images of the finite set of initial BS triplets:

$$((0,1),\epsilon,(1,2)),\quad ((0,1),\epsilon,(0,1)),\quad ((0,1),\epsilon,(0,2)),\quad ((1,2),\epsilon,(0,2)),$$
$$((1,2),\epsilon,(1,2)),\quad ((1,2),\epsilon,(0,1)),\quad ((0,2),\epsilon,(0,1)),\quad ((0,2),\epsilon,(0,2)),$$
$$((0,2),\epsilon,(1,2)),\quad ((1,2),1,(1,2)),\quad ((0,2),1,(1,2)),\quad ((0,2),1,(0,2)),$$
$$((1,2),0,(1,2)).$$

Example: intuitive definition of f-image

The morphism $\varphi = (0012, 2, 012)$:



All BS triplets are f^n-images of the finite set of initial BS triplets:

$$((0,01),\epsilon,(1,2)),\quad ((0,01),\epsilon,(0,2)),\quad ((0,012),\epsilon,(0,1)),\quad ((0,012),\epsilon,(0,2)),$$
$$((0,012),\epsilon,(1,2)),\quad ((0,22),\epsilon,(0,1)),\quad ((2,01),\epsilon,(0,2)),\quad ((0,012),0,(0,1)).$$

Structure of BS factors for non-repetitive D0L systems

Theorem (KK (2012))

For injective non-pushy and circular D0L system there exist well defined graphs of prolongations (finite number of vertices, each with out-degree equal to one) and a finite set of initial BS triplets such that each BS triplet is the f^n-image of an initial BS triplet.

Ingredients of the proof:

- vertices of the graphs are all pairs of w_1, w_2 such that
  - w_1, w_2 are factors of the language,
  - the last (resp. first) letters of w_1 and w_2 are distinct,
  - $|w_1| = |w_2| = C_S \cdot C_P$.
- Initial triplets $((w_1, w_3), v, (w_2, w_4))$ are those where v is non-synchronized.

Structure of BS factors for repetitive D0L systems (??)

Conjecture

For (injective) D0L system there exist well defined graphs of prolongations (finite number of vertices, each with out-degree equal to one) and a finite set of initial BS triplets such that each BS triplet is the f^n-image of an initial BS triplet.

Structure of BS factors for repetitive D0L systems (??)

Conjecture

For (injective) D0L system there exist well defined graphs of prolongations (finite number of vertices, each with out-degree equal to one) and a finite set of initial BS triplets such that each BS triplet is the f^n-image of an initial BS triplet.

What seems to be needed:

- Relaxing the definition of vertices for pushy D0L systems: (w_1, w_2) is a vertex with $|w_i|$ bounded by some constant or $w_i = u^\infty v$ where v is bounded and u is from a finite sets of factors.

Structure of BS factors for repetitive D0L systems (??)

Conjecture

For (injective) D0L system there exist well defined graphs of prolongations (finite number of vertices, each with out-degree equal to one) and a finite set of initial BS triplets such that each BS triplet is the f^n-image of an initial BS triplet.

What seems to be needed:

- Relaxing the definition of vertices for pushy D0L systems: (w_1, w_2) is a vertex with $|w_i|$ bounded by some constant or $w_i = u^\infty v$ where v is bounded and u is from a finite sets of factors.
- Redefine the set of initial BS triplets: the long enough non-synchronized BS triplets containing long repetitions must be treated separately.

Structure of BS factors for repetitive D0L systems (??)

Conjecture

For (injective) D0L system there exist well defined graphs of prolongations (finite number of vertices, each with out-degree equal to one) and a finite set of initial BS triplets such that each BS triplet is the f^n-image of an initial BS triplet.

What seems to be needed:

- Relaxing the definition of vertices for pushy D0L systems: (w_1, w_2) is a vertex with $|w_i|$ bounded by some constant or $w_i = u^\infty v$ where v is bounded and u is from a finite sets of factors.
- Redefine the set of initial BS triplets: the long enough non-synchronized BS triplets containing long repetitions must be treated separately.
- **We definitely need to know all infinite repetitions!**

Repetitive D0L systems

History:

1. (Ehrenfeucht, Rozenberg (1983)): Repetitiveness is decidable (but no reasonable algorithm).

Repetitive D0L systems

History:

1. (Ehrenfeucht, Rozenberg (1983)): Repetitiveness is decidable (but no reasonable algorithm).

2. (Mignosi, Séébold (1993)): Different proof of decidability (but no reasonable algorithm).

Repetitive D0L systems

History:

1. (Ehrenfeucht, Rozenberg (1983)): Repetitiveness is decidable (but no reasonable algorithm).
2. (Mignosi, Séébold (1993)): Different proof of decidability (but no reasonable algorithm).
3. (Kobayashi and Otto (2000)): Decidability and (sort of) reasonable algorithm.

Repetitive D0L systems

History:

1. (Ehrenfeucht, Rozenberg (1983)): Repetitiveness is decidable (but no reasonable algorithm).
2. (Mignosi, Séébold (1993)): Different proof of decidability (but no reasonable algorithm).
3. (Kobayashi and Otto (2000)): Decidability and (sort of) reasonable algorithm.

Related results:

- Question: given a morphism φ with an infinite fixed point $\mathbf{u}$ starting in a, is $\mathbf{u}$ eventually periodic?
  - (Pansiot (1986)), (Harju, Linna (1986)), (Honkala (2008))

Repetitive D0L systems

History:

1. (Ehrenfeucht, Rozenberg (1983)): Repetitiveness is decidable (but no reasonable algorithm).
2. (Mignosi, Séébold (1993)): Different proof of decidability (but no reasonable algorithm).
3. (Kobayashi and Otto (2000)): Decidability and (sort of) reasonable algorithm.

Related results:

- Question: given a morphism φ with an infinite fixed point $\mathbf{u}$ starting in a, is $\mathbf{u}$ eventually periodic?
  - (Pansiot (1986)), (Harju, Linna (1986)), (Honkala (2008))
- Question: given a morphism φ with an infinite fixed point $\mathbf{u}$ starting in a, is $\mathbf{u}$ purely periodic?
  - (Lando (1991)): simple algorithm is given (as we shall see ...)

Infinite periodic factors

Definition

Given a D0L system G, we say that v^∞ is an infinite periodic factor of G if v is a non-empty word and $v^k \in sub(L(G))$ for all positive integers k.
Let v be non-empty and primitive (not a power of shorter word). We say that infinite periodic factors v^∞ and u^∞ are equivalent if u is a power of a conjugate of v. We denote the equivalence class containing v^∞ by $[v]^\infty$.

Infinite periodic factors

Definition

Given a D0L system G, we say that v^∞ is an infinite periodic factor of G if v is a non-empty word and $v^k \in sub(L(G))$ for all positive integers k.
Let v be non-empty and primitive (not a power of shorter word). We say that infinite periodic factors v^∞ and u^∞ are equivalent if u is a power of a conjugate of v. We denote the equivalence class containing v^∞ by $[v]^\infty$.

Problem: Given a D0L system, find all $[v]^\infty$.

Infinite periodic factors over bounded letters

There is an infinite periodic factor $[v]^\infty$ in G iff G is pushy. We have already seen how to find all such v . . .

Infinite periodic factors containing an unbounded letter

Theorem (KK, ŠS (2013))

If $[v]^\infty$ is an infinite periodic factor of a D0L system $G = (\mathcal{A}, \varphi, w)$ such that $v \notin \mathcal{A}_0^+$, then there exist

- *u such that u^∞ is equivalent to v^∞,*
- *$a \in \mathcal{A}$ and $\ell \leq \#\mathcal{A}$ such that u^∞ is the fixed point of φ^ℓ starting with a.*

Infinite periodic factors containing an unbounded letter

Theorem (KK, ŠS (2013))

If $[v]^\infty$ is an infinite periodic factor of a D0L system $G = (\mathcal{A}, \varphi, w)$ such that $v \notin \mathcal{A}_0^+$, then there exist

- *u such that u^∞ is equivalent to v^∞,*
- *$a \in \mathcal{A}$ and $\ell \leq \#\mathcal{A}$ such that u^∞ is the fixed point of φ^ℓ starting with a.*

In other words: all infinite periodic factors containing an unbounded letter are purely periodic periodic points of φ.

But all purely periodic periodic points of an morphism can be found by the algorithm by Lando (1993)!

The algorithm by Lando

Problem: for a morphism φ over $\mathcal{A}$, letter $a \in \mathcal{A}$ and integer ℓ such that $\varphi^\ell(a) = av$ with $v \in \mathcal{A}^+$ decide whether $(\varphi^\ell)^\infty(a)$ is purely periodic:

1. If $v \in \mathcal{A}_0^+$, return the result: $(\varphi^\ell)^\infty(a)$ is not purely periodic (but eventually periodic).
2. Apply φ^ℓ to a until $(\varphi^\ell)^k(a)$ contains two occurrences of one unbounded letter.
3. If this letter is not a, then $(\varphi^\ell)^\infty(a)$ is not periodic, if it is, denote w the longest prefix containing a only as the first letter.
4. Now, $(\varphi^\ell)^\infty(a)$ is periodic if and only if $\varphi^\ell(w) = w^m$ for some integer $m \geq 2$.

Corollary: primitive morphisms generates circular languages

Theorem (Mossé (1991))

Let φ injective and primitive with an aperiodic fixed point $\mathbf{u}$, then the language of $\mathbf{u}$ is circular.

Důkaz.

We study D0L system $G = (\mathcal{A}, \varphi, a)$ where a is the first letter of $\mathbf{u}$.

- Clearly G is non-pushy (due to primitiveness).

Corollary: primitive morphisms generates circular languages

Theorem (Mossé (1991))

Let φ injective and primitive with an aperiodic fixed point $\mathbf{u}$, then the language of $\mathbf{u}$ is circular.

Důkaz.

We study D0L system $G = (\mathcal{A}, \varphi, a)$ where a is the first letter of $\mathbf{u}$.

- Clearly G is non-pushy (due to primitiveness).
- We know: repetitive iff circular and non-pushy (Mignosi, Séébold (1993)), therefore G is circular iff G is not repetitive.

Corollary: primitive morphisms generates circular languages

Theorem (Mossé (1991))

Let φ injective and primitive with an aperiodic fixed point $\mathbf{u}$, then the language of $\mathbf{u}$ is circular.

Důkaz.

We study D0L system $G = (\mathcal{A}, \varphi, a)$ where a is the first letter of $\mathbf{u}$.

- Clearly G is non-pushy (due to primitiveness).
- We know: repetitive iff circular and non-pushy (Mignosi, Séébold (1993)), therefore G is circular iff G is not repetitive.
- If G is repetitive, it must contain a purely periodic periodic point. A contradiction, since all periodic points have the same language!

□

Thank you for your attention! Thank you for your