Bifix codes and interval exchanges

Francesco Dolce

LIGM, 27^{th} January 2015

Joint work with :
V. Berthé, C. De Felice, J. Leroy, D. Perrin, C. Reutenauer and G. Rindone

Motivation

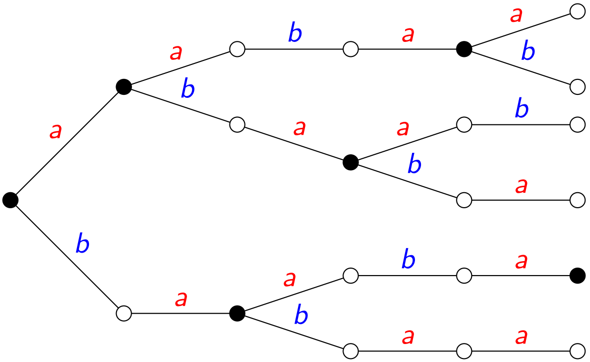
$$x = abaababaabaababa \cdots$$

$$x = \varphi^{\omega}(a)$$

$$\varphi : \begin{cases} a \mapsto ab \\ b \mapsto a \end{cases}$$

Motivation

$x = abaababaabaababa\cdots$

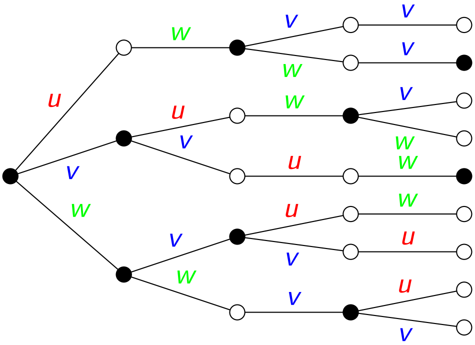


n	0	1	2	3	4	5	$\cdots$
(2-1)n+1	1	2	3	4	5	6	$\cdots$

Motivation

$x = \underline{ab}\,\underline{aa}\,\underline{ba}\,\underline{ba}\,\underline{ab}\,\underline{aa}\,\underline{ba}\,\underline{ba} \cdots$

$$\begin{cases} u = aa \\ v = ab \\ w = ba \end{cases}$$



n	0	1	2	3	4	$\cdots$
(3-1)n+1	1	3	5	7	9	$\cdots$

Motivation

$$x = v\,u\,w\,w\,v\,u\,w\,w\cdots$$

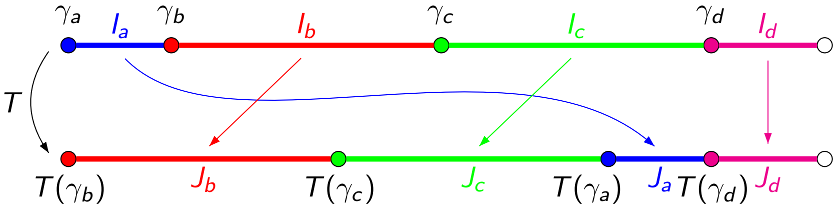
Outline

1. Interval Exchange Sets
2. Interval Exchange and Codes
3. Tree Sets

Outline

Let A be a finite set ordered by $<_1$ and $<_2$. An *interval exchange transformation* (IET) is a map $T : [0,1[\to [0,1[$ defined by

$$T(z) = z + \alpha_z \quad \text{if } z \in I_a.$$



$a <_1 b <_1 c <_1 d$ $b <_2 c <_2 a <_2 d$

A IET T is said to be *minimal* if for any $z \in [0, 1[$ the orbit $\mathcal{O}(z) = \{ T^n(z) \mid n \in \mathbb{Z}\}$ is dense in $[0, 1[$.

T is said *regular* if the orbits of the separation points $\neq 0$ are infinite and disjoint.

Theorem [Keane, 1975]

A regular interval exchange transformation is minimal.

Let T be an IET relative to $(I_a)_{a \in A}$. The *natural coding* of T relative to $z \in [0,1[$ is the infinite word $\Sigma_T(z) = a_0 a_1 \cdots \in A^\omega$ defined by

$$a_n = a \quad \text{si} \quad T^n(z) \in I_a.$$

Example

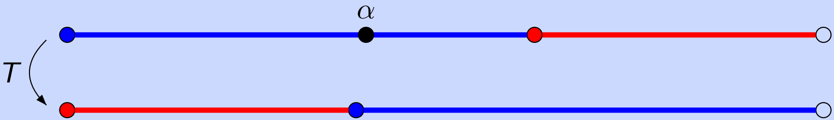
The *Fibonacci word* is the natural coding of the rotation of angle $\alpha = (3-\sqrt{5})/2$ relative to the point α, i.e. $T(z) = z + \alpha \bmod 1$.

Let T be an IET relative to $(I_a)_{a \in A}$. The *natural coding* of T relative to $z \in [0, 1[$ is the infinite word $\Sigma_T(z) = a_0 a_1 \cdots \in A^\omega$ defined by

$$a_n = a \quad \text{si} \quad T^n(z) \in I_a.$$

Example

The *Fibonacci word* is the natural coding of the rotation of angle $\alpha = (3 - \sqrt{5})/2$ relative to the point α, i.e. $T(z) = z + \alpha \bmod 1$.



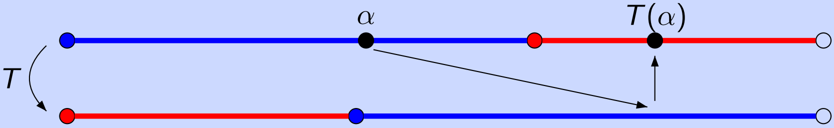
$\Sigma_T(\alpha) = a$

Let T be an IET relative to $(I_a)_{a\in A}$. The *natural coding* of T relative to $z \in [0,1[$ is the infinite word $\Sigma_T(z) = a_0 a_1 \cdots \in A^\omega$ defined by

$$a_n = a \quad \text{si } T^n(z) \in I_a.$$

Example

The *Fibonacci word* is the natural coding of the rotation of angle $\alpha = (3-\sqrt{5})/2$ relative to the point α, i.e. $T(z) = z + \alpha \bmod 1$.



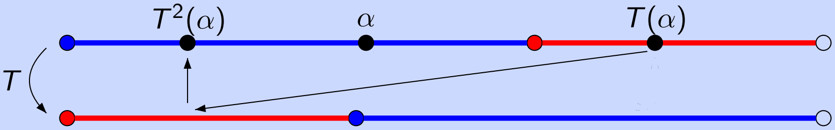
$$\Sigma_T(\alpha) = a\, b$$

Let T be an IET relative to $(I_a)_{a \in A}$. The *natural coding* of T relative to $z \in [0,1[$ is the infinite word $\Sigma_T(z) = a_0 a_1 \cdots \in A^\omega$ defined by

$$a_n = a \quad \text{si } T^n(z) \in I_a.$$

Example

The *Fibonacci word* is the natural coding of the rotation of angle $\alpha = (3-\sqrt{5})/2$ relative to the point α, i.e. $T(z) = z + \alpha \bmod 1$.



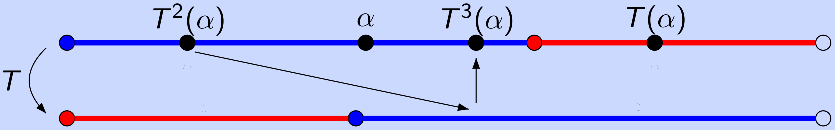
$$\Sigma_T(\alpha) = a\, b\, a$$

Let T be an IET relative to $(I_a)_{a\in A}$. The *natural coding* of T relative to $z \in [0,1[$ is the infinite word $\Sigma_T(z) = a_0a_1\cdots \in A^\omega$ defined by

$$a_n = a \quad \text{si} \quad T^n(z) \in I_a.$$

Example

The *Fibonacci word* is the natural coding of the rotation of angle $\alpha = (3-\sqrt{5})/2$ relative to the point α, i.e. $T(z) = z + \alpha \bmod 1$.



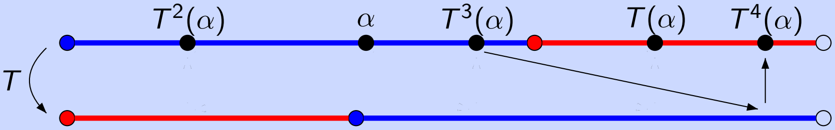
$$\Sigma_T(\alpha) = a\,b\,a\,a$$

Let T be an IET relative to $(I_a)_{a\in A}$. The *natural coding* of T relative to $z \in [0,1[$ is the infinite word $\Sigma_T(z) = a_0 a_1 \cdots \in A^\omega$ defined by

$$a_n = a \quad \text{si} \quad T^n(z) \in I_a.$$

Example

The *Fibonacci word* is the natural coding of the rotation of angle $\alpha = (3-\sqrt{5})/2$ relative to the point α, i.e. $T(z) = z + \alpha \bmod 1$.



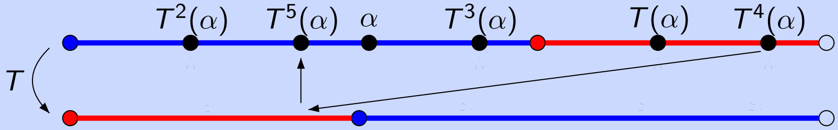
$$\Sigma_T(\alpha) = a\ b\ a\ a\ b$$

Let T be an IET relative to $(I_a)_{a\in A}$. The *natural coding* of T relative to $z \in [0,1[$ is the infinite word $\Sigma_T(z) = a_0 a_1 \cdots \in A^\omega$ defined by

$$a_n = a \quad \text{si } T^n(z) \in I_a.$$

Example

The *Fibonacci word* is the natural coding of the rotation of angle $\alpha = (3-\sqrt{5})/2$ relative to the point α, i.e. $T(z) = z + \alpha \bmod 1$.



$$\Sigma_T(\alpha) = a\, b\, a\, a\, b\, a \cdots$$

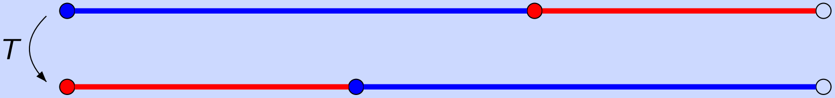
Proposition

If T is minimal, $F\big(\Sigma_T(z)\big)$ does not depend on z.

When T is regular (minimal), $F(T) = F\big(\Sigma_T(z)\big)$ is said a *regular* (*minimal*) *interval exchange set*.

Example

The *Fibonacci set* is the set of factors of a natural coding of the rotation of angle $\alpha = (3 - \sqrt{5})/2$.



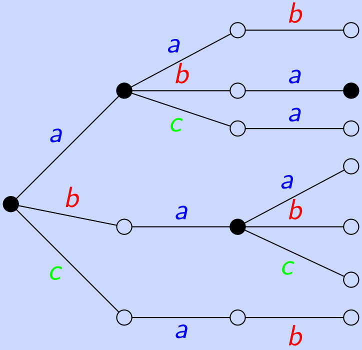
$$F(T) = \Big\{\varepsilon, a, b, aa, ab, ba, aab, aba, baa, \dots\Big\}$$

Proposition

Sturmian sets over an alphabet of cardinality $n > 2$ are not regular interval exchange sets.

Example

The *Tribonacci set* is a Sturmian set.



Outline

1. Interval Exchange Sets
2. Interval Exchange and Codes
   - Codes
   - Cylinders
   - Bifix Codes and Interval Exchanges
   - Coding Morphism
   - Maximal Bifix Decoding
3. Tree Sets

A set $X \subset A^+$ of nonempty words over an alphabet A is a *code* if for every $m, n \geq 1$ and $x_1, \ldots, x_n, y_1, \ldots, y_m$,

$$x_1 \cdots x_n = y_1 \cdots y_m \implies n = m \text{ and } x_i = y_i \text{ for } i = 1, \ldots, n.$$

A *prefix code* is a set of nonempty words which does not contain any proper prefix of its elements. A *suffix code* is defined symmetrically. A *bifix code* is a set which is both a prefix code and a suffix code.

Example

- $\{a, ab, ba\}$ is not a code.
- $\{aabb, ababb, abb\}$ is a prefix code but it's not a suffix code.
- $\{aa, ab, ba\}$ is a bifix code.

A set $X \subset A^+$ of nonempty words over an alphabet A is a *code* if for every $m, n \geq 1$ and $x_1, \ldots, x_n, y_1, \ldots, y_m$,

$$x_1 \cdots x_n = y_1 \cdots y_m \quad \Longrightarrow \quad n = m \quad \text{and} \quad x_i = y_i \text{ for } i = 1, \ldots, n.$$

A *prefix code* is a set of nonempty words which does not contain any proper prefix of its elements. A *suffix code* is defined symmetrically. A *bifix code* is a set which is both a prefix code and a suffix code.

Example

- $\{a, ab, ba\}$ is not a code.
- $\{aabb, ababb, abb\}$ is a prefix code but it's not a suffix code.
- $\{aa, ab, ba\}$ is a bifix code.

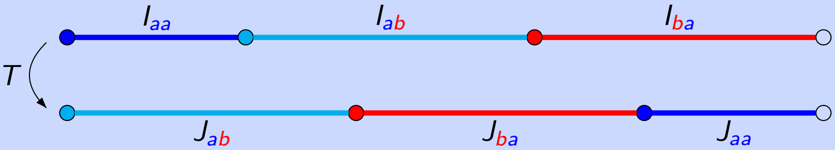
A bifix code $X \subset S$ is S-*maximal* if it is not properly contained in a bifix code $Y \subset S$.

For a word $w = b_0 b_1 \cdots b_{m-1}$, let's define

$$I_w = I_{b_0} \cap T^{-1}(I_{b_1}) \cap \ldots \cap T^{-m+1}(I_{b_{m-1}})$$

and $J_w = T^m(I_w)$.

Example



$I_{aa} = I_a \cap T^{-1}(I_a)$, $I_{ab} = I_a \cap T^{-1}(I_b)$, $I_{ba} = I_b \cap T^{-1}(I_a)$, $I_{bb} = I_b \cap T^{-1}(I_b)$;

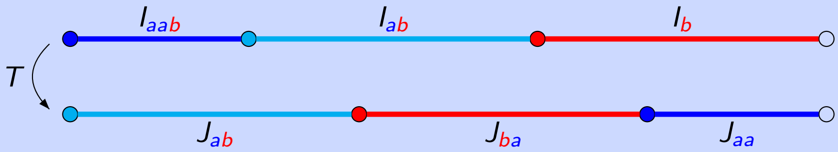
$J_{aa} = T^2(I_a) \cap T(I_a)$, $J_{ab} = T^2(I_a) \cap T(I_b)$, $J_{ba} = T^2(I_b) \cap T(I_a)$, $J_{bb} = T^2(I_b) \cap T(I_b)$.

We denote by $<_1$ the lexicographic order on A^* induced by $<_1$ and by $<_2$ the lexicographic order on the reversal of the words induced by $<_2$

Proposition

- $I_u < I_v$ if and only if $u <_1 v$ and u is not a prefix of v.
- $J_u < J_v$ if and only if $u <_2 v$ and u is not a suffix of v.

Example



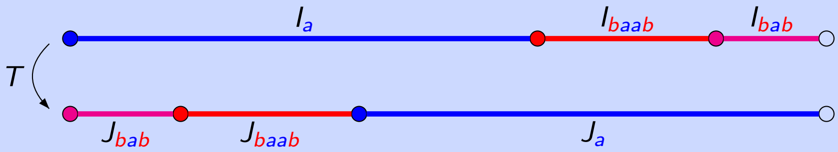
$aab <_1 ab <_1 b$ while $ab <_2 ba <_2 aa$.

Proposition

Let T a minimal IET. If X is a finite $F(T)$-maximal bifix code, the families $(I_w)_{w\in X}$ and $(J_w)_{w\in X}$ are ordered partitions of $[0,1[$, relatively to the orders $<_1$ and $<_2$.

Example

Let S be the Fibonacci set. The set $X = \{a, baab, bab\}$ is an S-maximal bifix code.

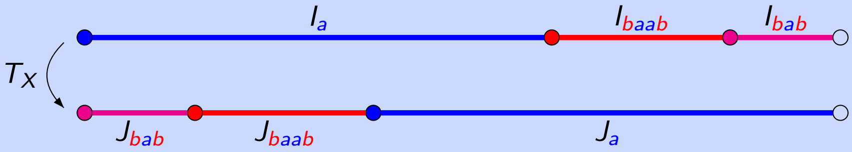


$a <_1 baab <_1 bab$ and $bab <_2 baab <_2 a$.

Let T be a regular IET. Let X be a finite $F(T)$-maximal bifix code on the alphabet A. Let's define the transformation

$$T_X(z) = T^{|u|}(z) \quad \text{if } z \in I_u.$$

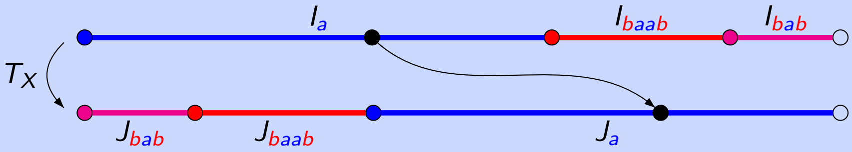
Example



Let T be a regular IET. Let X be a finite $F(T)$-maximal bifix code on the alphabet A. Let's define the transformation

$$T_X(z) = T^{|u|}(z) \quad \text{if } z \in I_u.$$

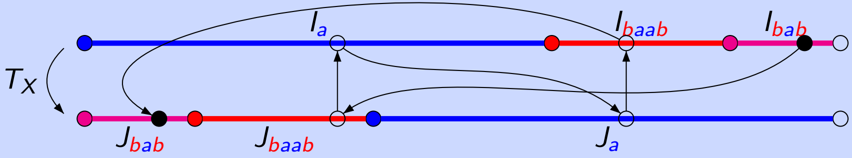
Example



Let T be a regular IET. Let X be a finite $F(T)$-maximal bifix code on the alphabet A. Let's define the transformation

$$T_X(z) = T^{|u|}(z) \quad \text{if } z \in I_u.$$

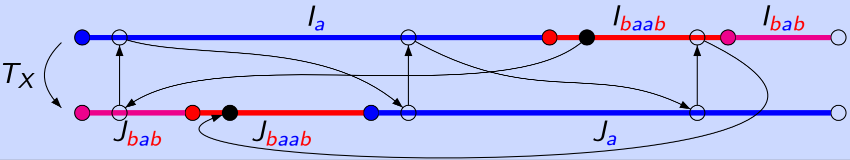
Example



Let T be a regular IET. Let X be a finite $F(T)$-maximal bifix code on the alphabet A. Let's define the transformation

$$T_X(z) = T^{|u|}(z) \quad \text{if } z \in I_u.$$

Example



A *coding morphism* for a prefix code $X \subset A^+$ is a morphism $f : B^* \to A^*$ which maps bijectively B onto X.

Example

Let's consider the bifix code $X = \{aa, ab, ba\}$ on $A = \{a, b\}$ and let $B = \{u, v, w\}$.

The map

$$f : \begin{cases} u \mapsto aa \\ v \mapsto ab \\ w \mapsto ba \end{cases}$$

is a coding morphism for X.

Let $f : B^* \to A^*$ be a coding morphism for X. Let $(K_b)_{b \in B}$, with $K_b = I_{f(b)}$. Let T_f be the IET relative to $(K_b)_{b \in B}$.

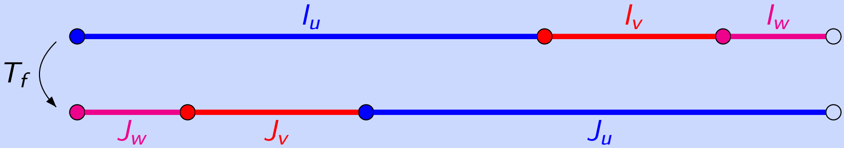
Proposition

If X is a finite $F(T)$-maximal bifix code, one has $T_f = T_X$.

Example

Let $X = \{a, baab, bab\}$, $B = \{u, v, w\}$ and

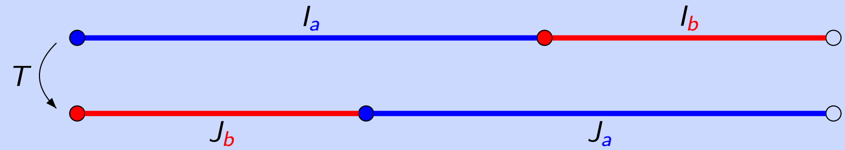
$$f : u \mapsto a, \quad v \mapsto baab, \quad w \mapsto bab.$$



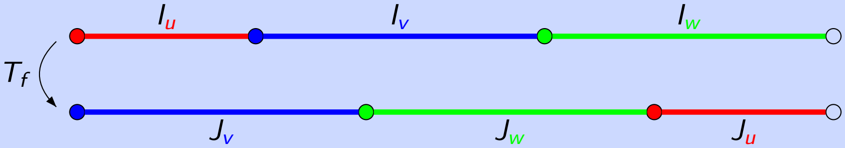
Theorem [2014]

Let T a regular IET. For any finite $F(T)$-maximal bifix code X with coding morphism f, the transformation T_f is regular.

Example



$X = \{aa, ab, ba\}$ and $f : u \mapsto aa, v \mapsto ab, w \mapsto ba$.

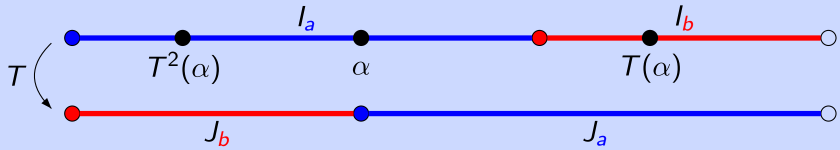


Let f be a coding morphism for a S-maximal prefix code. The *decoding* of x is the infinite word y s.t. $x = f(y)$.

Proposition

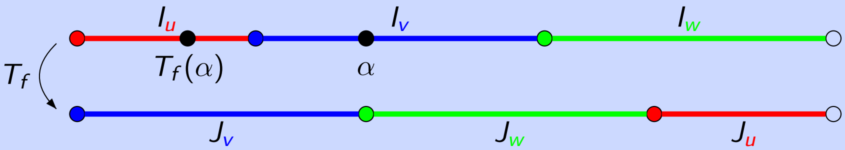
Let T be a minimal IET, X a finite $F(T)$-maximal prefix code and $f : B^* \to A^*$ a coding morphism.
Then, for all $z \in [0, 1[$, one has $\Sigma_T(z) = f(\Sigma_{T_f}(z))$.

Example



$$\Sigma_T(\alpha) = ab\,aa\,ba\,ba\,ab\,aa\cdots$$

$X = \{aa, ab, ba\}$ and $f : u \mapsto aa, v \mapsto ab, w \mapsto ba$.



$$\Sigma_{T_f}(\alpha) = v\,u\,w\,w\,v\,u\cdots$$

$$f\,(vuwwvu\cdots) = \underline{ab}\,\underline{aa}\,\underline{ba}\,\underline{ba}\,\underline{ab}\,\underline{aa}\,\underline{ba}\,\underline{ba}\cdots$$

Let f be a coding morphism for a finite S-maximal bifix code $X \subset S$. The set $f^{-1}(S)$ is called a *maximal bifix decoding* of S.

Theorem [2014]

The family of regular interval exchange sets is closed under maximal bifix decoding.

Proof. $f^{-1}(S) = F(T_f)$.

Actually, this property is true for a larger class of sets...

Outline

Let S be a biextendable set of words. For $w \in S$, we denote

$$L(w) = \{a \in A \mid aw \in S\}, \qquad R(w) = \{a \in A \mid wa \in S\}$$

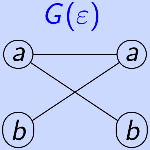
and

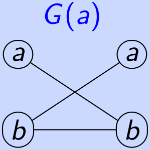
$$E(w) = \{(a, b) \in A \times A \mid awb \in S\}.$$

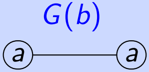
The *extension graph* of w is the undirected bipartite graph $G(w)$ with vertices $L(w) \sqcup R(w)$ and edges $E(w)$.

Example

Let S be the Fibonacci set.





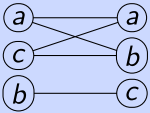


We say that a biextendable set S is a *tree set* if the graph $G(w)$ is a tree (connected and acyclic) for all $w \in S$.

Example

Let $A = \{a, b, c\}$. The set S of factors of $a^*\{bc, bcbc\}a^*$ is not a tree set.

$G(\varepsilon)$

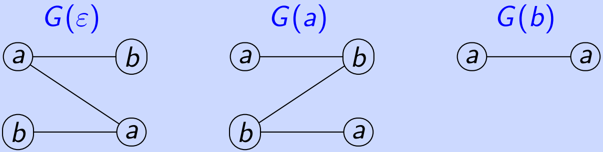


Let $<_1$ and $<_2$ be two orders on A. For a set S and a word $w \in S$, the graph $G(w)$ is *compatible* with $<_1$ and $<_2$ if for any $(a,b),(c,d) \in E(w)$, one has

$$a <_1 b \quad \Longrightarrow \quad b \leq_2 d$$

Example

Let S be the Fibonacci set. Set $a <_1 b$ and $b <_2 a$.

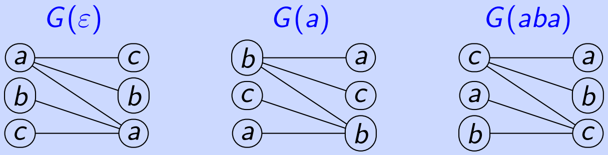


We say that a biextendable set S is a *planar tree set* w.r.t. $<_1$ and $<_2$ on A if for any $w \in S$, the graph $G(w)$ is a tree compatible with $<_1$ and $<_2$.

Example

Let $A = \{a, b, c\}$. The *Tribonacci set* is the set of factors of the Tribonacci word, i.e. is the fixpoint $x = f^\omega(a) = abacaba\cdots$ of the morphism

$$f : a \mapsto ab, \quad b \mapsto ac, \quad c \mapsto a.$$

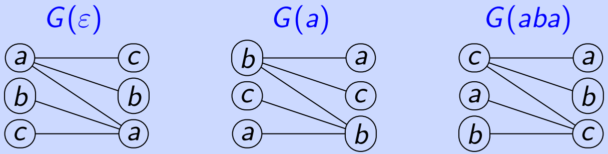


It is not possible to find two orders on A making the three graphs planar.

Example

Let $A = \{a, b, c\}$. The *Tribonacci set* is the set of factors of the Tribonacci word, i.e. is the fixpoint $x = f^\omega(a) = abacaba\cdots$ of the morphism

$$f : a \mapsto ab, \quad b \mapsto ac, \quad c \mapsto a.$$



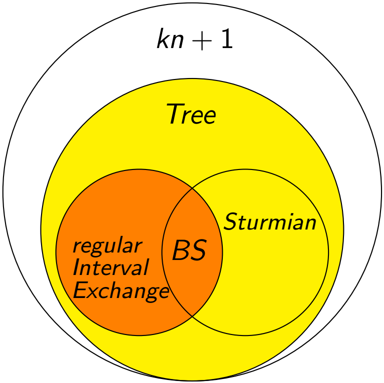
It is not possible to find two orders on A making the three graphs planar.

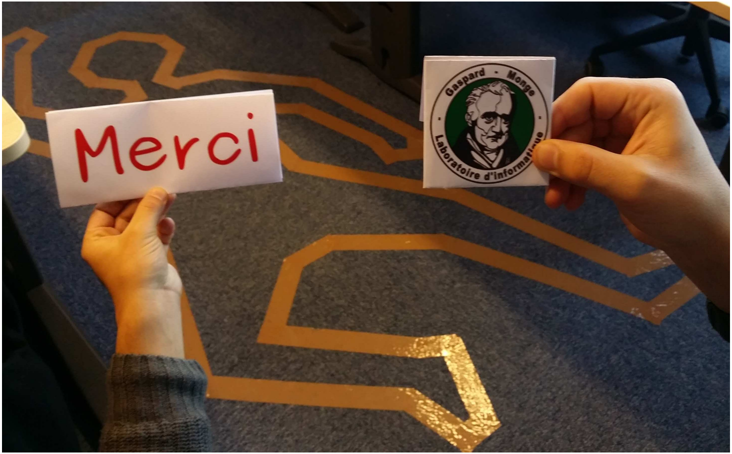
Theorem [Ferenczi, Zamboni, 2008]

A set S is a regular interval exchange set on A if and only if it is a uniformly recurrent planar tree set containing A.

Theorem [2014]

The family of uniformly recurrent tree set is closed under maximal bifix decoding.




Merci
Gaspard · Monge
Laboratoire d'informatique