Clustering of return words in languages of Interval Exchange Transformations

Francesco Dolce

joint work with Christian B. Hughes

LIGM - Université Gustave Eiffel

Marne-la-Vallée, 4 avril 2025

louche – chelou

louche – chelou

fou – ouf

louche – chelou

fou – ouf

Two words uv and vu are *conjugates*.

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

`b a n a n a`

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

b	a	n	a	n	a					
	a	n	a	n	a	b				
		n	a	n	a	b	a			
			a	n	a	b	a	n		
				n	a	b	a	n	a	
					a	b	a	n	a	n

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

a	b	a	n	a	n
a	n	a	b	a	n
a	n	a	n	a	b
b	a	n	a	n	a
n	a	b	a	n	a
n	a	n	a	b	a

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

a	b	a	n	a	n
a	n	a	b	a	n
a	n	a	n	a	b
b	a	n	a	n	a
n	a	b	a	n	a
n	a	n	a	b	a

$\mathrm{bwt}_{\mathcal{A}}(\mathtt{banana}) = \mathtt{nnbaaa}$

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

a_1	b	a	n	a	n_1
a_2	n	a	b	a	n_2
a_3	n	a	n	a	b_1
b_1	a	n	a	n	a_1
n_1	a	b	a	n	a_2
n_2	a	n	a	b	a_3

$$\mathrm{bwt}_{\mathcal{A}}(\mathtt{banana}) = \mathtt{nnbaaa}$$

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

n
n
b
a
a
a

$$\mathrm{bwt}_{\mathcal{A}}(\mathtt{banana}) = \mathtt{nnbaaa}$$

$$\mathrm{bwt}_{\mathcal{A}}^{-1}(\mathtt{nnbaaa}) = ?$$

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

a	n
a	n
a	b
b	a
n	a
n	a

$$\mathrm{bwt}_{\mathcal{A}}(\mathtt{banana}) = \mathtt{nnbaaa}$$

$$\mathrm{bwt}^{-1}_{\mathcal{A}}(\mathtt{nnbaaa}) = ?$$

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

a		n
a		n
a		b
b		a
n	a	a
n		a

$$\mathrm{bwt}_{\mathcal{A}}(\mathtt{banana}) = \mathtt{nnbaaa}$$

$$\mathrm{bwt}_{\mathcal{A}}^{-1}(\mathtt{nnbaaa}) = ?$$

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

a		n
a		n
a		b
b		a
n	a	a
n	a	a

$\mathrm{bwt}_{\mathcal{A}}\,(\mathtt{banana}) = \mathtt{nnbaaa}$

$\mathrm{bwt}_{\mathcal{A}}^{-1}\,(\mathtt{nnbaaa}) = ?$

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

a		n
a		n
a		b
b	a	a
n	a	a
n	a	a

$$\mathrm{bwt}_{\mathcal{A}}\,(\mathtt{banana}) = \mathtt{nnbaaa}$$

$$\mathrm{bwt}_{\mathcal{A}}^{-1}\,(\mathtt{nnbaaa}) = ?$$

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

a	b	n
a		n
a		b
b	a	a
n	a	a
n	a	a

$\mathrm{bwt}_{\mathcal{A}}(\mathtt{banana}) = \mathtt{nnbaaa}$

$\mathrm{bwt}_{\mathcal{A}}^{-1}(\mathtt{nnbaaa}) = ?$

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

a	b	n
a	n	n
a		b
b	a	a
n	a	a
n	a	a

$\text{bwt}_{\mathcal{A}}(\texttt{banana}) = \texttt{nnbaaa}$

$\text{bwt}_{\mathcal{A}}^{-1}(\texttt{nnbaaa}) = ?$

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

a	b	n
a	n	n
a	n	b
b	a	a
n	a	a
n	a	a

$$\mathrm{bwt}_{\mathcal{A}}(\texttt{banana}) = \texttt{nnbaaa}$$

$$\mathrm{bwt}_{\mathcal{A}}^{-1}(\texttt{nnbaaa}) = ?$$

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

a	b	a	n	a	n
a	n	a	b	a	n
a	n	a	n	a	b
b	a	n	a	n	a
n	a	b	a	n	a
n	a	n	a	b	a

$$\mathrm{bwt}_{\mathcal{A}}(\mathtt{banana}) = \mathtt{nnbaaa}$$

$$\mathrm{bwt}_{\mathcal{A}}^{-1}(\mathtt{nnbaaa}) = ?$$

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

a	b	a	n	a	n
a	n	a	b	a	n
a	n	a	n	a	b
b	a	n	a	n	a
n	a	b	a	n	a
n	a	n	a	b	a

$$\mathrm{bwt}_{\mathcal{A}}(\mathtt{banana}) = \mathtt{nnbaaa}$$

$$\mathrm{bwt}^{-1}_{\mathcal{A}}(\mathtt{nnbaaa}) = [\mathtt{banana}]$$

Burrows-Wheeler Transform

Let $\mathcal{A} = \{a_1 < a_2 < \cdots < a_d\}$ and $w \in \mathcal{A}^*$.

a	b	a	n	a	n
a	n	a	b	a	n
a	n	a	n	a	b
b	a	n	a	n	a
n	a	b	a	n	a
n	a	n	a	b	a

$$\mathrm{bwt}_{\mathcal{A}}(\mathtt{banana}) = \mathtt{nnbaaa}$$

$$\mathrm{bwt}^{-1}_{\mathcal{A}}(\mathtt{nnbaaa}) = [\mathtt{banana}] = [\mathtt{abanan}]$$

Burrows-Wheeler Transform

Proposition

$[w] = [u]$ if and only if $\mathrm{bwt}_{\mathcal{A}}(u) = \mathrm{bwt}_{\mathcal{A}}(w)$

Burrows-Wheeler Transform

Proposition

$[w] = [u]$ if and only if $\mathrm{bwt}_{\mathcal{A}}(u) = \mathrm{bwt}_{\mathcal{A}}(w)$

Proposition

$[w] = [u^p]$ if and only if $\mathrm{bwt}_{\mathcal{A}}(u) = b_1 \cdots b_n$ and $\mathrm{bwt}_{\mathcal{A}}(w) = b_1^p \cdots b_n^p$.

$\mathrm{bwt}_{\mathcal{A}}(\texttt{bon}) = \texttt{nob}$ $\qquad$ $\mathrm{bwt}_{\mathcal{A}}(\texttt{bonbon}) = \texttt{nnoobb}$

CTU CZECH TECHNICAL UNIVERSITY IN PRAGUE

Burrows-Wheeler Transform

Proposition

$[w] = [u]$ if and only if $\mathrm{bwt}_{\mathcal{A}}(u) = \mathrm{bwt}_{\mathcal{A}}(w)$

Proposition

$[w] = [u^p]$ if and only if $\mathrm{bwt}_{\mathcal{A}}(u) = b_1 \cdots b_n$ and $\mathrm{bwt}_{\mathcal{A}}(w) = b_1^p \cdots b_n^p$.

$$\mathrm{bwt}_{\mathcal{A}}(\mathtt{bon}) = \mathtt{nob} \qquad \mathrm{bwt}_{\mathcal{A}}(\mathtt{bonbon}) = \mathtt{nnoobb}$$

Proposition

$\mathrm{bwt}_{\mathcal{A}}(\cdot)$ gives an injection between the set of Lyndon words and $\mathcal{A}^*$.

w is *Lyndon* if primitive and smallest in $[w]$.

Burrows-Wheeler Transform

Proposition

$[w] = [u]$ if and only if $\mathrm{bwt}_{\mathcal{A}}(u) = \mathrm{bwt}_{\mathcal{A}}(w)$

Proposition

$[w] = [u^p]$ if and only if $\mathrm{bwt}_{\mathcal{A}}(u) = b_1 \cdots b_n$ and $\mathrm{bwt}_{\mathcal{A}}(w) = b_1^p \cdots b_n^p$.

$\mathrm{bwt}_{\mathcal{A}}(\mathtt{bon}) = \mathtt{nob} \qquad \mathrm{bwt}_{\mathcal{A}}(\mathtt{bonbon}) = \mathtt{nnoobb}$

Proposition

$\mathrm{bwt}_{\mathcal{A}}(\cdot)$ gives an injection between the set of Lyndon words and $\mathcal{A}^*$.

w is *Lyndon* if primitive and smallest in $[w]$.

$\mathrm{bwt}_{\mathcal{A}}^{-1}(\mathtt{babaaa}) = ?$

Extendend Burrows-Wheeler Transform

Let W be the multiset $\{\mathtt{aab}, \mathtt{ab}, \mathtt{ab}\}$ of Lyndon words over $\{\mathtt{a} <_\omega \mathtt{b}\}$.

a	a	b
a	b	a
a	b	
a	b	
b	a	a
b	a	
b	a	

Extendend Burrows-Wheeler Transform

Let W be the multiset $\{\texttt{aab}, \texttt{ab}, \texttt{ab}\}$ of Lyndon words over $\{\texttt{a} <_\omega \texttt{b}\}$.

a	a	b
a	b	a
a	b	
a	b	
b	a	a
b	a	
b	a	

$$\mathrm{ebwt}_{\mathcal{A}}(W) = \texttt{babbaaa}$$

Extendend Burrows-Wheeler Transform

Let W be the multiset $\{\mathtt{aab}, \mathtt{ab}, \mathtt{ab}\}$ of Lyndon words over $\{\mathtt{a} <_\omega \mathtt{b}\}$.

a	a	b
a	b	a
a	b	
a	b	
b	a	a
b	a	
b	a	

$$\mathrm{ebwt}_{\mathcal{A}}(W) = \mathtt{babbaaa}$$

Theorem [Mantaci, Restivo, Rosone, Sciortino (2007)]

$\mathrm{ebwt}_{\mathcal{A}}(\cdot)$ gives a bijection between the family of multisets of Lyndon words and $\mathcal{A}^*$.

Clustering words

A word $w \in \mathcal{A}^*$ is π-*clustering* for $\mathcal{A}$ if $\mathrm{bwt}_{\mathcal{A}}(w) = a_{\pi(1)}^{k_1} \cdots a_{\pi(d)}^{k_d}$, where $k_i = |w|_{a_{\pi(i)}}$.

Clustering words

A word $w \in \mathcal{A}^*$ is π-*clustering* for $\mathcal{A}$ if $\mathrm{bwt}_{\mathcal{A}}(w) = a_{\pi(1)}^{k_1} \cdots a_{\pi(d)}^{k_d}$, where $k_i = |w|_{a_{\pi(i)}}$.
It is *perfectly clustering* if π is the symmetric permutation $\pi(i) = d - i + 1$.

Clustering words

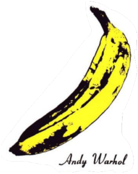
A word $w \in \mathcal{A}^*$ is π-*clustering* for $\mathcal{A}$ if $\mathrm{bwt}_{\mathcal{A}}(w) = a_{\pi(1)}^{k_1} \cdots a_{\pi(d)}^{k_d}$, where $k_i = |w|_{a_{\pi(i)}}$.
It is *perfectly clustering* if π is the symmetric permutation $\pi(i) = d - i + 1$.

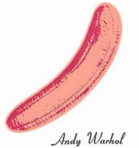
$$\mathcal{A} = \{\mathtt{a} < \mathtt{b} < \mathtt{n}\}, \qquad \mathcal{A}' = \{\mathtt{a} < \mathtt{n} < \mathtt{b}\}, \qquad \mathcal{A}'' = \{\mathtt{n} < \mathtt{a} < \mathtt{b}\}$$

$\mathrm{bwt}_{\mathcal{A}}(\mathtt{banana}) = \mathtt{nnbaaa}$, $\pi = (\mathtt{n}\ \mathtt{b}\ \mathtt{a})$

$\mathrm{bwt}_{\mathcal{A}'}(\mathtt{banana}) = \mathtt{bnnaaa}$, $\pi' = (\mathtt{b}\ \mathtt{n}\ \mathtt{a})$

$\mathrm{bwt}_{\mathcal{A}''}(\mathtt{banana}) = \mathtt{aabnna}$ not clustering





Clustering words

Proposition

Let $w = u^p \in \mathcal{A}^*$. Then w is π-clustering for $\mathcal{A}$ if and only if u is π-clustering for $\mathcal{A}$.

Clustering words

Proposition

Let $w = u^p \in \mathcal{A}^*$. Then w is π-clustering for $\mathcal{A}$ if and only if u is π-clustering for $\mathcal{A}$.

Theorem [Mantaci, Restivo, Sciortino (2003)]

Over a binary alphabet, (perfectly) clustering words are exactly powers of Christoffel words (i.e., finite standard Sturmian words) and their conjugates.

Clustering words

Proposition

Let $w = u^p \in \mathcal{A}^*$. Then w is π-clustering for $\mathcal{A}$ if and only if u is π-clustering for $\mathcal{A}$.

Theorem [Mantaci, Restivo, Sciortino (2003)]

Over a binary alphabet, (perfectly) clustering words are exactly powers of Christoffel words (i.e., finite standard Sturmian words) and their conjugates.

And over larger alphabets ?

Clustering words and (D)IETs

Theorem [Ferenczi, Zamboni (2013)]

Let $w \in \mathcal{A}^*$ be primitive, with $\mathrm{Card}\,(\mathcal{A}) = d$. The following are equivalent :

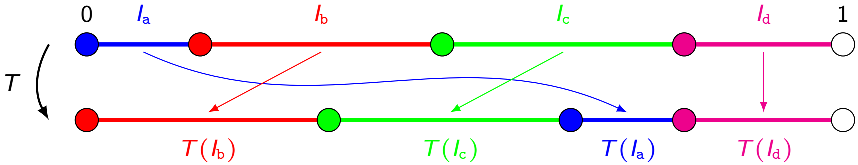
1. w is π-clustering ;
2. ww occurs in the trajectory of a minimal d-DIET with permutation π ;
3. ww occurs in the trajectory of a d-DIET with permutation π ;
4. ww occurs in the trajectory of a d-IET with permutation π.

IETs ? DIETs ?

Interval exchanges

Let $(I_a)_{a\in\mathcal{A}}$ be a partition of $[\ell, r)$.
An *interval exchange transformation* (IET) is a map $T : [\ell, r) \to [\ell, r)$ defined by

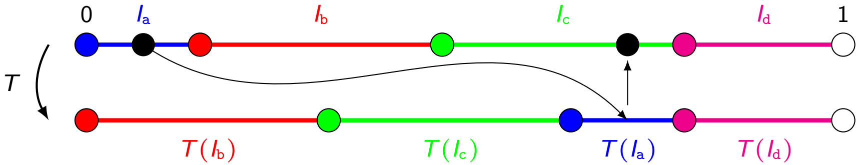
$$T(z) = z + \tau_a \quad \text{if} \quad z \in I_a.$$



Interval exchanges

Let $(I_a)_{a \in \mathcal{A}}$ be a partition of $[\ell, r)$.
An *interval exchange transformation* (IET) is a map $T : [\ell, r) \to [\ell, r)$ defined by

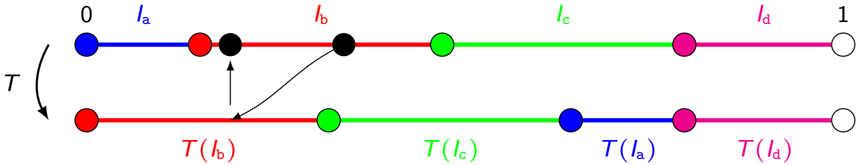
$$T(z) = z + \tau_a \quad \text{if} \quad z \in I_a.$$



Interval exchanges

Let $(I_a)_{a \in \mathcal{A}}$ be a partition of $[\ell, r)$.
An *interval exchange transformation* (IET) is a map $T : [\ell, r) \to [\ell, r)$ defined by

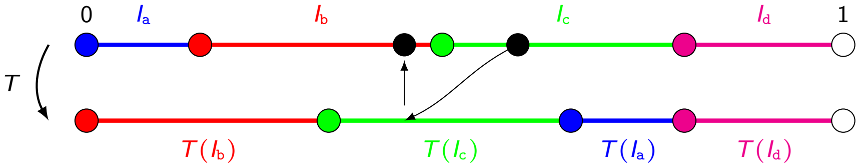
$$T(z) = z + \tau_a \quad \text{if} \quad z \in I_a.$$



Interval exchanges

Let $(I_a)_{a \in \mathcal{A}}$ be a partition of $[\ell, r)$.
An *interval exchange transformation* (IET) is a map $T : [\ell, r) \to [\ell, r)$ defined by

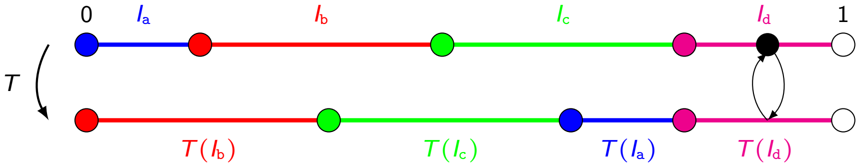
$$T(z) = z + \tau_a \quad \text{if} \quad z \in I_a.$$



Interval exchanges

Let $(I_a)_{a\in\mathcal{A}}$ be a partition of $[\ell, r)$.
An *interval exchange transformation* (IET) is a map $T : [\ell, r) \to [\ell, r)$ defined by

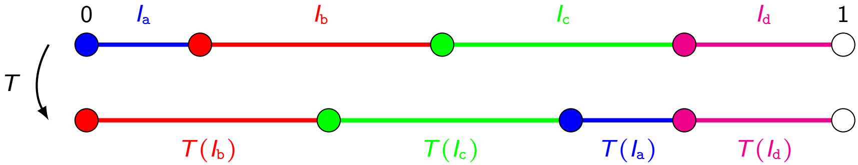
$$T(z) = z + \tau_a \quad \text{if} \quad z \in I_a.$$



Interval exchanges

Let $(I_a)_{a \in \mathcal{A}}$ be a partition of $[\ell, r)$.
An *interval exchange transformation* (IET) is a map $T : [\ell, r) \to [\ell, r)$ defined by

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



$$\mathcal{A} = \{a < b < c < d\} \qquad \pi = (b\ c\ a\ d)$$

Minimality and regularity

T is *minimal* if for any $z \in [\ell, r)$ the orbit $\mathcal{O}(z) = \{T^n(z) \mid n \in \mathbb{Z}\}$ is dense in $[\ell, r)$.

T is *regular* if the orbits of the non-zero formal discontinuities are infinite and disjoint.

Theorem [Keane (1975)]

A regular interval exchange transformation is minimal.

Minimality and regularity

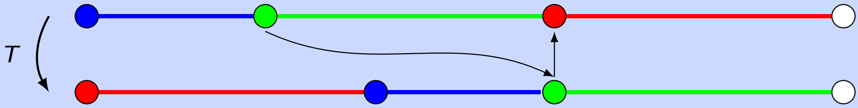
T is *minimal* if for any $z \in [\ell, r)$ the orbit $\mathcal{O}(z) = \{T^n(z) \mid n \in \mathbb{Z}\}$ is dense in $[\ell, r)$.

T is *regular* if the orbits of the non-zero formal discontinuities are infinite and disjoint.

Theorem [Keane (1975)]

A regular interval exchange transformation is minimal.

Example (the converse is not true)

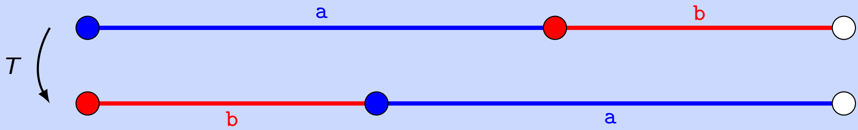


Trajectories

The *trajectory* of $z \in [\ell, r)$ under T is the infinite word $\Omega_T(z) = w_0 w_1 \cdots \in \mathcal{A}^\omega$ defined by

$$w_n = a \qquad \text{if } T^n(z) \in I_a.$$

Example (Fibonacci)

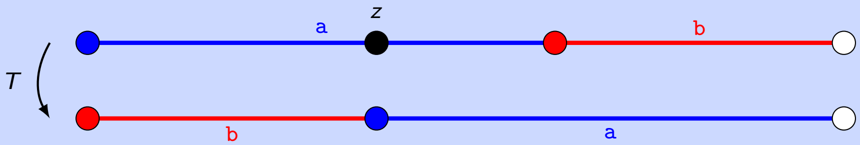


Trajectories

The *trajectory* of $z \in [\ell, r)$ under T is the infinite word $\Omega_T(z) = w_0 w_1 \cdots \in \mathcal{A}^\omega$ defined by

$$w_n = a \qquad \text{if } T^n(z) \in I_a.$$

Example (Fibonacci)



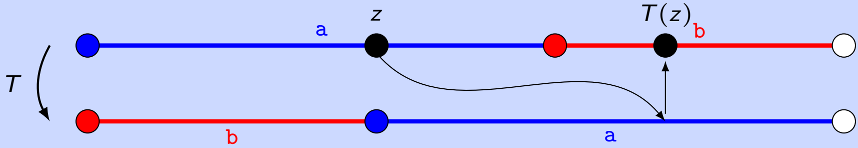
$\Omega_T(z) = \mathtt{a}$

Trajectories

The *trajectory* of $z \in [\ell, r)$ under T is the infinite word $\Omega_T(z) = w_0 w_1 \cdots \in \mathcal{A}^\omega$ defined by

$$w_n = a \qquad \text{if } T^n(z) \in I_a.$$

Example (Fibonacci)



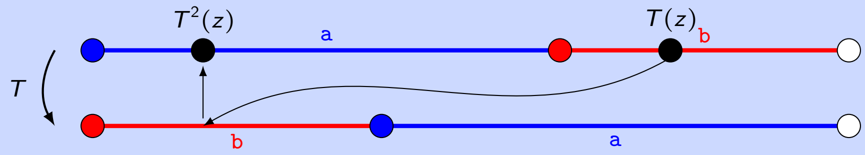
$$\Omega_T(z) = ab$$

Trajectories

The *trajectory* of $z \in [\ell, r)$ under T is the infinite word $\Omega_T(z) = w_0 w_1 \cdots \in \mathcal{A}^\omega$ defined by

$$w_n = a \qquad \text{if } T^n(z) \in I_a.$$

Example (Fibonacci)



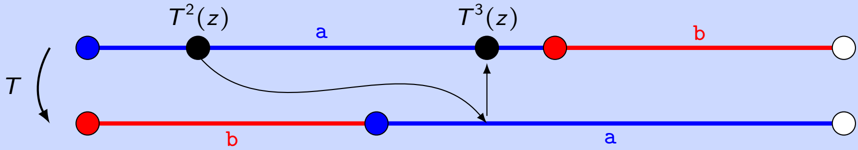
$\Omega_T(z) = \mathtt{aba}$

Trajectories

The *trajectory* of $z \in [\ell, r)$ under T is the infinite word $\Omega_T(z) = w_0 w_1 \cdots \in \mathcal{A}^\omega$ defined by

$$w_n = a \qquad \text{if } T^n(z) \in I_a.$$

Example (Fibonacci)



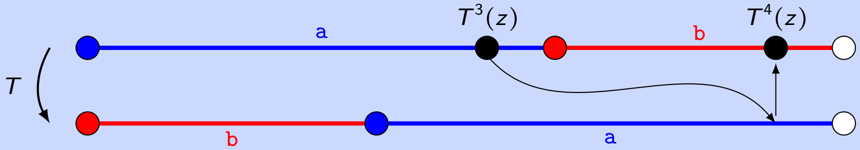
$$\Omega_T(z) = \mathtt{abaa}$$

Trajectories

The *trajectory* of $z \in [\ell, r)$ under T is the infinite word $\Omega_T(z) = w_0 w_1 \cdots \in \mathcal{A}^\omega$ defined by

$$w_n = a \qquad \text{if } T^n(z) \in I_a.$$

Example (Fibonacci)



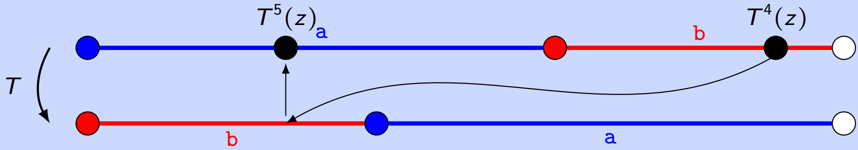
$\Omega_T(z) = \mathtt{abaab}$

Trajectories

The *trajectory* of $z \in [\ell, r)$ under T is the infinite word $\Omega_T(z) = w_0 w_1 \cdots \in \mathcal{A}^\omega$ defined by

$$w_n = a \qquad \text{if } T^n(z) \in I_a.$$

Example (Fibonacci)

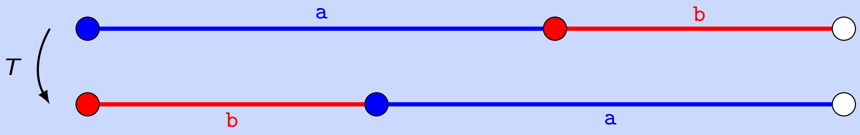


$\Omega_T(z) = \mathtt{abaaba}\cdots$

Interval exchange languages

The set $\mathcal{L}(T) = \bigcup_{z \in [\ell, r)} \mathcal{L}(\Omega_T(z))$ is a (*minimal*, *regular*) *interval exchange language*.

Example (Fibonacci)

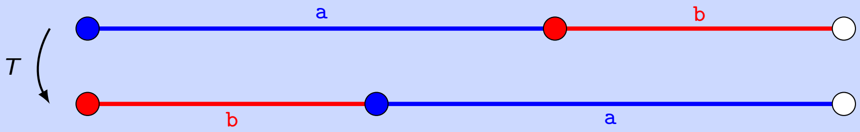


Interval exchange languages

The set $\mathcal{L}(T) = \bigcup_{z \in [\ell, r)} \mathcal{L}(\Omega_T(z))$ is a (*minimal*, *regular*) *interval exchange language*.

Remark. When T is minimal, $\mathcal{L}(\Omega_T(z))$ does not depend on the point z.

Example (Fibonacci)

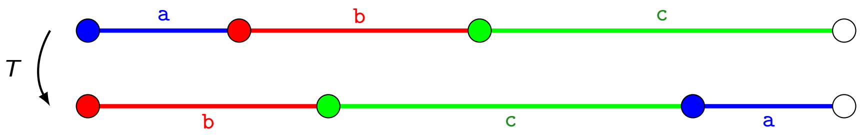


$$\mathcal{L}(T) = \{\varepsilon, \mathtt{a}, \mathtt{b}, \mathtt{aa}, \mathtt{ab}, \mathtt{ba}, \mathtt{aab}, \mathtt{aba}, \mathtt{baa}, \mathtt{bab}, \mathtt{aaba}, \ldots\}$$

From letters to words

Given a IET T and a word $w = w_0 w_1 \cdots w_m \in \mathcal{A}^*$, let

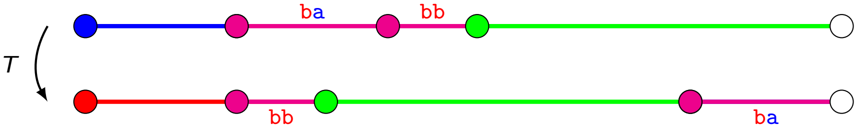
$$I_w = I_{w_0} \cap T^{-1}(I_{w_1}) \cap \cdots \cap T^{-m}(I_{w_m}) \qquad \text{(by convention } I_\varepsilon = [\ell, r))$$



From letters to words

Given a IET T and a word $w = w_0 w_1 \cdots w_m \in \mathcal{A}^*$, let

$$I_w = I_{w_0} \cap T^{-1}(I_{w_1}) \cap \cdots \cap T^{-m}(I_{w_m}) \qquad \text{(by convention } I_\varepsilon = [\ell, r)\text{)}$$



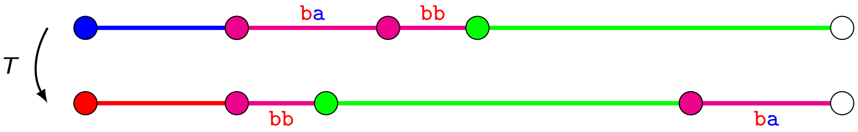
$$I_{ba} = I_b \cap T^{-1}(I_a) \qquad T^2(I_{ba}) = T^2(I_b) \cap T(I_a)$$

Thus $\Omega_T(z)$ starts with w for every $z \in I_w$.

From letters to words

Given a IET T and a word $w = w_0 w_1 \cdots w_m \in \mathcal{A}^*$, let

$$I_w = I_{w_0} \cap T^{-1}(I_{w_1}) \cap \cdots \cap T^{-m}(I_{w_m}) \qquad \text{(by convention } I_\varepsilon = [\ell, r)\text{)}$$



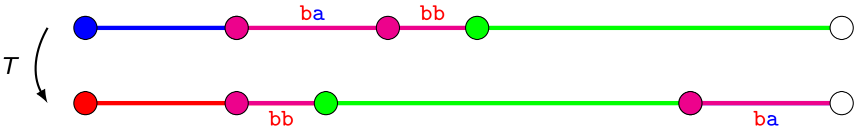
Proposition

If T is minimal, then $w \in \mathcal{L}(T) \Leftrightarrow |I_w| \neq 0$.

From letters to words

Given a IET T and a word $w = w_0 w_1 \cdots w_m \in \mathcal{A}^*$, let

$$I_w = I_{w_0} \cap T^{-1}(I_{w_1}) \cap \cdots \cap T^{-m}(I_{w_m}) \qquad \text{(by convention } I_\varepsilon = [\ell, r))$$



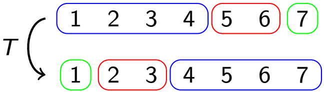
Proposition

- I_u is on the left of I_v if and only if $u <_{\mathcal{A}} v$ and u is not a prefix of v.
- $T^{|u|}(I_u)$ is on the left of $T^{|v|}(I_v)$ if and only if $\tilde{u} <_\pi \tilde{v}$ and u is not a suffix of v.

From IETs to DIETs

Let $(n_1, n_2, \ldots, n_d)$ be a composition of $n = \sum n_i$.
A *discrete interval exchange transformation* (DIET) is a map $T : \mathbb{N}_n \to \mathbb{N}_n$ defined by

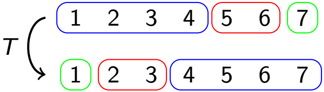
$$T(k) = k + t_i \quad \text{if } \; k \in \left[\sum_{j<i} n_j, \; \sum_{j \leq i} n_j \right].$$



From IETs to DIETs

Let $(n_1, n_2, \ldots, n_d)$ be a composition of $n = \sum n_i$.
A *discrete interval exchange transformation* (DIET) is a map $T : \mathbb{N}_n \to \mathbb{N}_n$ defined by

$$T(k) = k + t_i \quad \text{if} \quad k \in \left[\sum_{j<i} n_j, \sum_{j \le i} n_j \right].$$

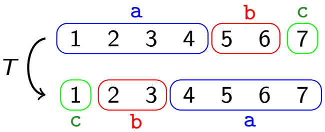


$$\Omega_T(1) = 1\,4\,7\,1\,4\,7\,1\,4\,7 \cdots = (1\,4\,7)^\omega$$

From IETs to DIETs

Let $(n_1, n_2, \ldots, n_d)$ be a composition of $n = \sum n_i$.
A *discrete interval exchange transformation* (DIET) is a map $T : \mathbb{N}_n \to \mathbb{N}_n$ defined by

$$T(k) = k + t_i \quad \text{if} \quad k \in \left[\sum_{j<i} n_j, \; \sum_{j \leq i} n_j \right].$$



$$\Omega_T(1) = \mathtt{a\,a\,c\,a\,a\,c\,a\,a\,c} \cdots = (\mathtt{a\,a\,c})^\omega$$

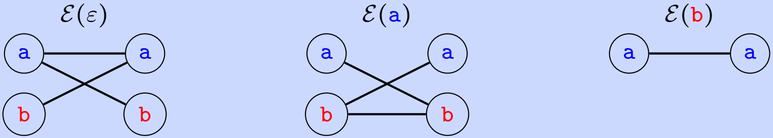
Extension graphs

The *extension graph* of $w \in \mathcal{L}$ is the bipartite graph $\mathcal{E}(w)$ with vertices

$$L(w) = \{a \in \mathcal{A} \mid aw \in \mathcal{L}\} \qquad \text{and} \qquad R(w) = \{b \in \mathcal{A} \mid wb \in \mathcal{L}\},$$

and edges $B(w) = \{(a, b) \in \mathcal{A} \times \mathcal{A} \mid awb \in \mathcal{L}\}$.

Example (Fibonacci, $\mathcal{L} = \{\varepsilon, \mathrm{a}, \mathrm{b}, \mathrm{aa}, \mathrm{ab}, \mathrm{ba}, \mathrm{aba}, \mathrm{baa}, \mathrm{bab}, \ldots\}$)



Extension graphs

The *extension graph* of $w \in \mathcal{L}$ is the bipartite graph $\mathcal{E}(w)$ with vertices

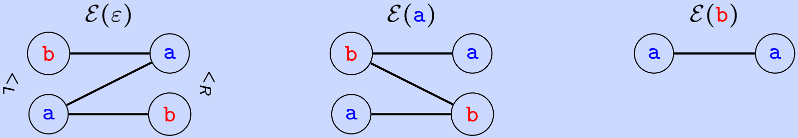
$$L(w) = \{a \in \mathcal{A} \mid aw \in \mathcal{L}\} \qquad \text{and} \qquad R(w) = \{b \in \mathcal{A} \mid wb \in \mathcal{L}\},$$

and edges $B(w) = \{(a,b) \in \mathcal{A} \times \mathcal{A} \mid awb \in \mathcal{L}\}$.

$\mathcal{E}(w)$ is *compatible* with two orders $<_L, <_R$ on $\mathcal{A}$ if for every $(a,b),(c,d) \in B(w)$

$$a <_L c \quad \Rightarrow \quad b \leq_R d.$$

Example (Fibonacci, $\mathtt{b} <_L \mathtt{a}$, and $\mathtt{a} <_R \mathtt{b}$)



See the forest for the IETs

It's all Greek to me!

$\mathcal{L}$ is *dendric* (δένδρον) if every $\mathcal{E}(w)$ is a tree. It is *alsinic* (ἄλσος) $\mathcal{E}(w)$ is a forest.

See the forest for the IETs

It's all Greek to me!

$\mathcal{L}$ is *dendric* (δένδρον) if every $\mathcal{E}(w)$ is a tree. It is *alsinic* (ἄλσος) $\mathcal{E}(w)$ is a forest.

$\mathcal{L}$ is *ordered dendric* for $<_L$ and $<_R$ if every $\mathcal{E}(w)$ is compatible with the orders.

See the forest for the IETs

It's all Greek to me !

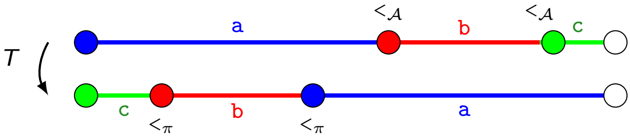
$\mathcal{L}$ is *dendric* (δένδρον) if every $\mathcal{E}(w)$ is a tree. It is *alsinic* (ἄλσος) $\mathcal{E}(w)$ is a forest.

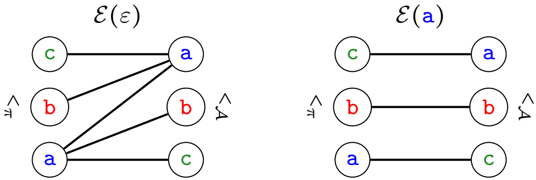
$\mathcal{L}$ is *ordered dendric* for $<_L$ and $<_R$ if every $\mathcal{E}(w)$ is compatible with the orders.

Theorem [Ferenczi, Zamboni (2008) ; Ferenczi, Hubert, Zamboni (2024)]

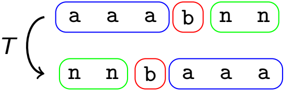
- $\mathcal{L}$ is an interval exchange language iff it is (uniformly) recurrent ordered alsinic.
- $\mathcal{L}$ is a minimal interval exchange language iff it is aperiodic (uniformly) recurrent ordered alsinic.
- $\mathcal{L}$ is a regular interval exchange language iff it is (uniformly) recurrent ordered dendric.

See the forest for the IETs



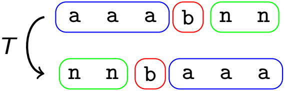


See the forest for the DIETs



$$\mathrm{bwt}_{\{a<b<n\}}(\mathtt{banana}) = \mathtt{nnbaaa}$$

See the forest for the DIETs

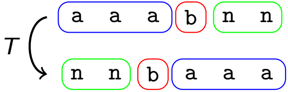


$\text{bwt}_{\{a<b<n\}}(\text{banana}) = \text{nnbaaa}$

Theorem [Ferenczi, Hubert, Zamboni (2023)]

A word $w \in \mathcal{A}^*$ is π-clustering if and only if for every bispecial word $v \in \mathcal{L}(w^\omega)$ the graph $\mathcal{E}(v)$ is compatible with $<_\mathcal{A}$ and $<_\pi$.

See the forest for the DIETs



$$\mathrm{bwt}_{\{a<b<n\}}(\mathtt{banana}) = \mathtt{nnbaaa}$$

Theorem [Ferenczi, Hubert, Zamboni (2023)]

A word $w \in \mathcal{A}^*$ is π-clustering if and only if for every bispecial word $v \in \mathcal{L}(w^\omega)$ the graph $\mathcal{E}(v)$ is compatible with $<_\mathcal{A}$ and $<_\pi$.

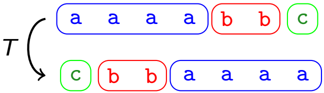
Corollary

A word $w \in \mathcal{A}^*$ is π-clustering if and only if $\mathcal{L}(w^\omega)$ is ordered alsinic for $<_\mathcal{A}$ and $<_\pi$.

…and for multisets

Proposition

If a multiset $W \subset \mathcal{A}^*$ is π-clustering, then every $w \in W$ is π_w-clustering, with π_w the restriction of π to the letters of w.

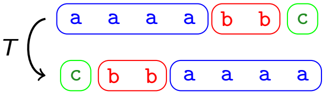


$\mathrm{ebwt}_{\{a<b<c\}}(\{aac, ab, ab\}) = cbbaaaa$; $\quad \mathrm{bwt}_{\{a<c\}}(aac) = caa$, $\quad \mathrm{bwt}_{\{a<b\}}(ab) = ba$

...and for multisets

Proposition

If a multiset $W \subset \mathcal{A}^*$ is π-clustering, then every $w \in W$ is π_w-clustering, with π_w the restriction of π to the letters of w.



$\mathrm{ebwt}_{\{a<b<c\}}(\{aac, ab, ab\}) = cbbaaaa$; $\mathrm{bwt}_{\{a<c\}}(aac) = caa$, $\mathrm{bwt}_{\{a<b\}}(ab) = ba$

Example (the converse is not true)

$\mathrm{bwt}_{\{a<b\}}(ab) = ba$, $\mathrm{bwt}_{\{a<b\}}(aab) = baa$, but $\mathrm{ebwt}_{\{a<b\}}(\{ab, aab\}) = babaa$.

Return words

Νόστοι

Question [Lapointe (2021)]

Are return words of a symmetric IET perfectly clustering words ?

Return words

Νόστοι

Question [Lapointe (2021)]

Are return words of a symmetric IET perfectly clustering words ?

$$\mathcal{R}_{\mathcal{L}}(w) = \{u \in \mathcal{A}^* \mid wu \in (\mathcal{L} \cap \mathcal{A}^* w) \setminus \mathcal{A}^+ w \mathcal{A}^+\}$$

Example (Fibonacci)

$$\mathbf{f} = \underline{aba}\,\underline{aba}\,baaba\,\underline{aba}\,\underline{ba}\,ababaaba\cdots$$

$\mathcal{R}(\mathtt{aba}) = \{\mathtt{aba}, \mathtt{ba}\}$

Induced transformations

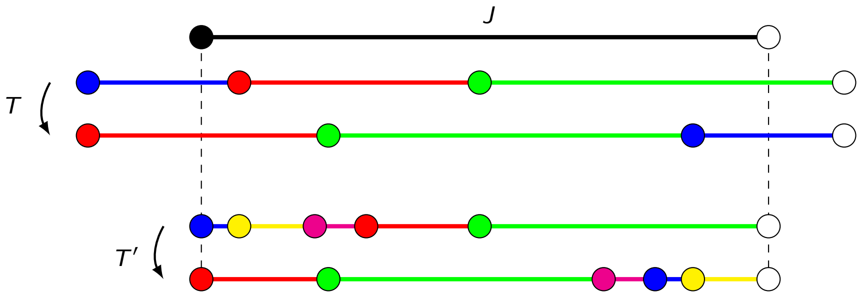
Let T be a minimal IET and $J \subset [\ell, r)$.
The *transformation induced* by T (*first return map* of T) on J is $T' : J \to J$ defined by

$$T'(z) = T^{\nu(z)}(z) \qquad \text{with} \;\; \nu(z) = \min\{n > 0 \mid T^n(z) \in J\}$$

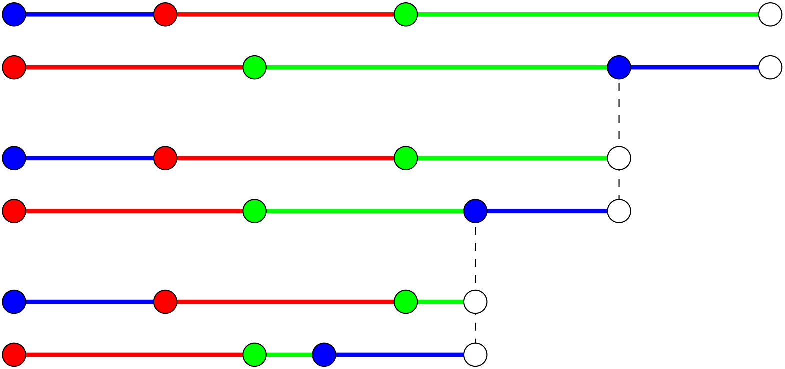
Induced transformations

Let T be a minimal IET and $J \subset [\ell, r)$.
The *transformation induced* by T (*first return map* of T) on J is $T' : J \to J$ defined by

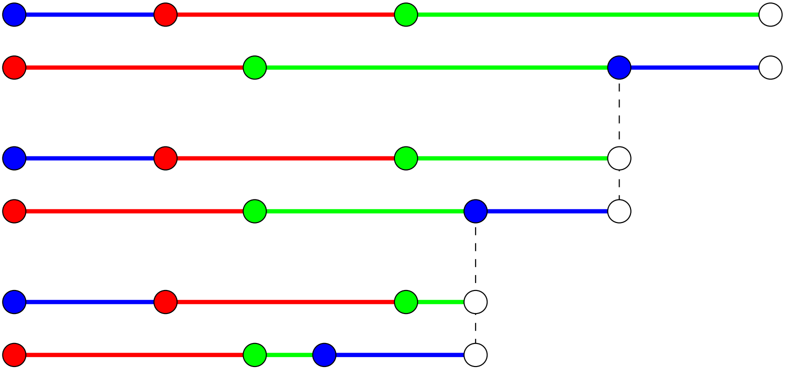
$$T'(z) = T^{\nu(z)}(z) \qquad \text{with } \nu(z) = \min\{n > 0 \mid T^n(z) \in J\}$$



Right Rauzy induction



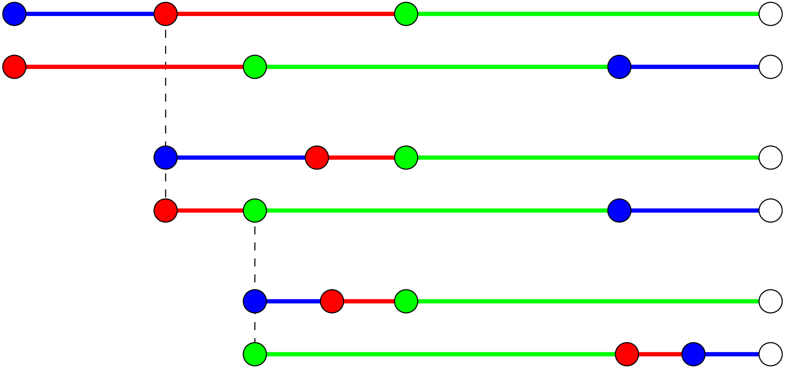
Right Rauzy induction



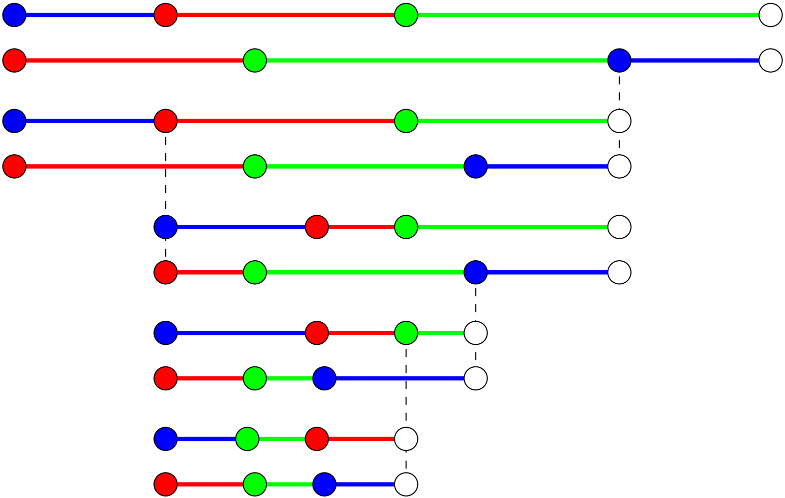
Theorem [Rauzy (1979)]

If T is regular, the right Rauzy induced transformation is regular on the same alphabet.

Left Rauzy induction



Two-sided Rauzy induction



Rauzy induction on I_w

Theorem [D., Perrin (2017)]

Let T be a regular IET and $w \in \mathcal{L}(T)$.
The regular IET T' induced on I_w can be obtained by two-sided Rauzy induction.

Rauzy induction on I_w

Theorem [D., Perrin (2017)]

Let T be a regular IET and $w \in \mathcal{L}(T)$.
The regular IET T' induced on I_w can be obtained by two-sided Rauzy induction.

Moreover, each step is associated with an automorphism of $F_{\mathcal{A}}$ of the form

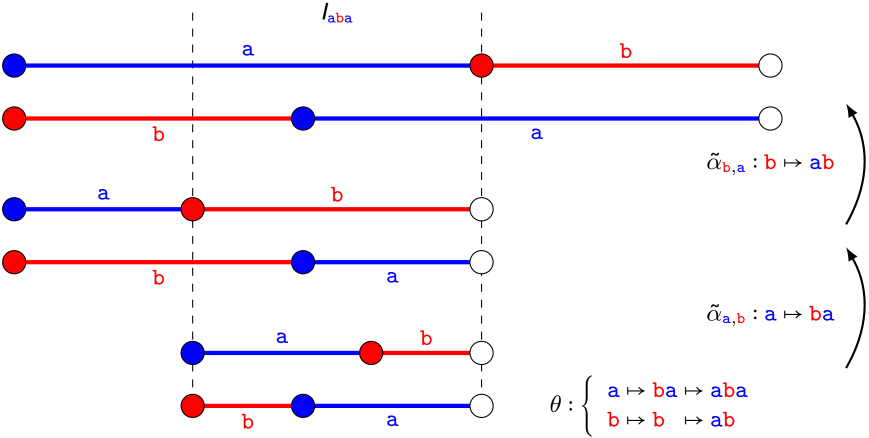
$$\alpha_{a,b} = \begin{cases} a \mapsto ab \\ c \mapsto c \end{cases} \qquad \text{or} \qquad \tilde{\alpha}_{a,b} = \begin{cases} a \mapsto ba \\ c \mapsto c \end{cases}$$

and for every $z \in I_w$

$$\Omega_T(z) = \theta\left(\Omega_{T'}(z)\right).$$

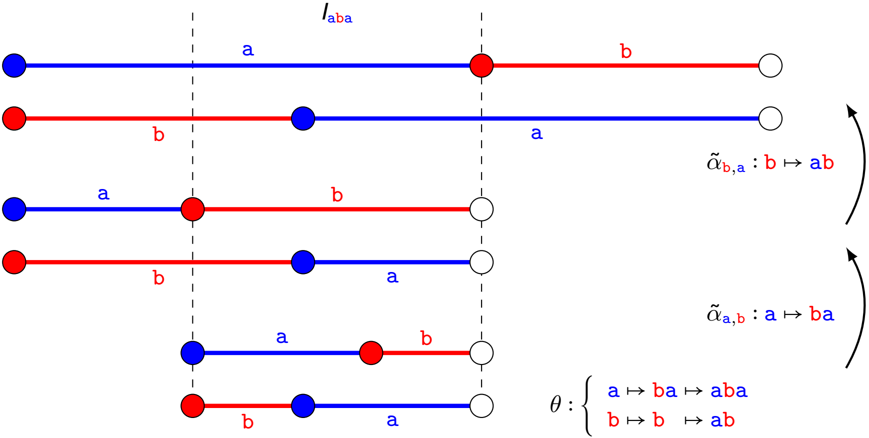
(θ is obtained from the morphisms above, but backwards!)

Rauzy induction on I_w



$$\theta : \begin{cases} a \mapsto ba \mapsto aba \\ b \mapsto b \mapsto ab \end{cases}$$

Rauzy induction on I_w



$$\theta : \begin{cases} a \mapsto ba \mapsto aba \\ b \mapsto b \mapsto ab \end{cases}$$

$$\mathcal{R}(aba) = \{aba, ab\}$$

Return words in IETs are clustering

a.k.a. the main result

Corollary

$\mathcal{R}(w) = \theta(\mathcal{A})$.

Return words in IETs are clustering

a.k.a. the main result

Corollary

$\mathcal{R}(w) = \theta(\mathcal{A})$.

Trivial fact

Letters are clustering.

Return words in IETs are clustering

a.k.a. the main result

Corollary

$\mathcal{R}(w) = \theta(\mathcal{A})$.

Trivial fact

Letters are clustering.

Proposition (way too technical to be stated properly)

Under the "right conditions" each $\alpha_{a,b}$ (resp., $\tilde{\alpha}_{a,b}$) preserves clustering.

Return words in IETs are clustering

a.k.a. the main result

Corollary

$\mathcal{R}(w) = \theta(\mathcal{A})$.

Trivial fact

Letters are clustering.

Proposition (way too technical to be stated properly)

Under the "right conditions" each $\alpha_{a,b}$ (resp., $\tilde{\alpha}_{a,b}$) preserves clustering.

Theorem

Let T be regular IET and $w \in \mathcal{L}(T)$. Each $u \in \mathcal{R}(w)$ is clustering.

Where to go from here ?

- What about the permutation π ?

Where to go from here ?

- What about the permutation π ?

- What if T is a non minimal (D)IET ?

Where to go from here ?

- What about the permutation π ?

- What if T is a non minimal (D)IET ?

$$\begin{matrix} a & a & a & a & b & b & c \\ c & b & b & a & a & a & a \end{matrix}$$

$$\mathcal{L}\left((aac)^\omega \cup (ab)^\omega \cup (ab)^\omega\right)$$

$$\begin{matrix} a & a & a & c & b & b \\ c & b & b & a & a & a \end{matrix}$$

$$\theta : \begin{cases} a \mapsto a \mapsto a \\ b \mapsto ab \mapsto ab \\ c \mapsto c \mapsto ac \end{cases}$$

$$\begin{matrix} a & b & b & c \\ c & b & b & a \end{matrix}$$

$$\mathcal{R}(a) = \{a,\ ab,\ ac\}$$

Cimer coupbeau !