Dendric languages and the Finite Index Basis Property

Francesco Dolce

CTU CZECH TECHNICAL UNIVERSITY IN PRAGUE

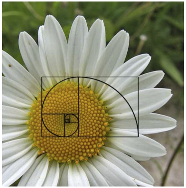
Conference on Theoretical and Computational Algebra

Pocinho, 6 de julho de 2023

Fibonacci

$$\mathbf{x} = \mathtt{abaababaabaababa}\cdots$$

$$\mathbf{x} = \lim_{\mathbf{n}\to\infty} \varphi^{\mathbf{n}}(\mathtt{a}) \qquad \text{where} \qquad \varphi : \begin{cases} \mathtt{a} \mapsto \mathtt{ab} \\ \mathtt{b} \mapsto \mathtt{a} \end{cases}$$



Fibonacci

$\mathbf{x} = \text{abaababaabaababa} \cdots$

The Fibonacci word is a *Sturmian word*. Its set of factor $\mathcal{L}(\mathbf{x})$ is a Sturmian language.

Definition

A *Sturmian* language $\mathcal{L} \subset \mathcal{A}^*$ is a factorial set such that $p_n = \text{Card}\,(\mathcal{L} \cap \mathcal{A}^n) = n + 1$.

$$\mathcal{L}(\mathbf{x}) = \{\underbrace{\varepsilon}_{1}, \underbrace{\text{a}, \text{b}}_{2}, \underbrace{\text{aa}, \text{ab}, \text{ba}}_{3}, \underbrace{\text{aab}, \text{aba}, \text{baa}, \text{bab}}_{4}, \underbrace{\text{aaba}, \text{abaa}, \text{abab}, \text{baab}, \text{baba}}_{5}, \ldots\}$$

2-coded Fibonacci

$\mathbf{x} = \texttt{ab aa ba ba ab aa ba ba} \cdots$

2-coded Fibonacci

$$\mathbf{x} = \texttt{ab aa ba ba ab aa ba ba} \cdots$$

$$f^{-1}(\mathbf{x}) = \texttt{v u w w v u w w} \cdots$$

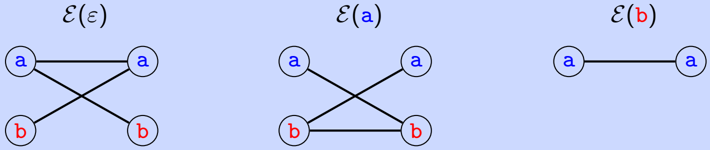
$$f : \begin{cases} \texttt{u} & \mapsto & \texttt{aa} \\ \texttt{v} & \mapsto & \texttt{ab} \\ \texttt{w} & \mapsto & \texttt{ba} \end{cases}$$

Extension graphs

The *extension graph* of a word $w \in \mathcal{L}$ is the undirected bipartite graph $\mathcal{E}(w)$ with vertices $L(w) \sqcup R(w)$ and edges $B(w)$, where

$$\begin{aligned} L(w) &= \{u \in \mathcal{A} \mid uw \in \mathcal{L}\} \\ R(w) &= \{v \in \mathcal{A} \mid wv \in \mathcal{L}\} \\ B(w) &= \{(u,v) \in \mathcal{A} \times \mathcal{A} \mid uwv \in \mathcal{L}\} \end{aligned}$$

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

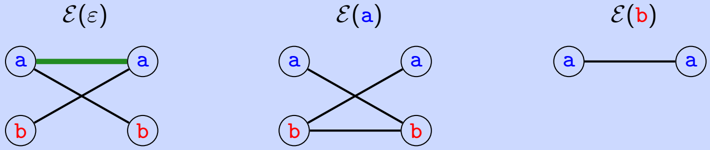


Extension graphs

The *extension graph* of a word $w \in \mathcal{L}$ is the undirected bipartite graph $\mathcal{E}(w)$ with vertices $L(w) \sqcup R(w)$ and edges $B(w)$, where

$$\begin{aligned} L(w) &= \{u \in \mathcal{A} \mid uw \in \mathcal{L}\} \\ R(w) &= \{v \in \mathcal{A} \mid wv \in \mathcal{L}\} \\ B(w) &= \{(u,v) \in \mathcal{A} \times \mathcal{A} \mid uwv \in \mathcal{L}\} \end{aligned}$$

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

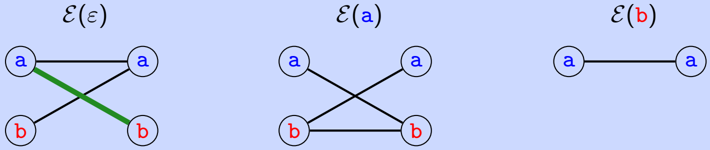


Extension graphs

The *extension graph* of a word $w \in \mathcal{L}$ is the undirected bipartite graph $\mathcal{E}(w)$ with vertices $L(w) \sqcup R(w)$ and edges $B(w)$, where

$$\begin{aligned} L(w) &= \{u \in \mathcal{A} \mid uw \in \mathcal{L}\} \\ R(w) &= \{v \in \mathcal{A} \mid wv \in \mathcal{L}\} \\ B(w) &= \{(u,v) \in \mathcal{A} \times \mathcal{A} \mid uwv \in \mathcal{L}\} \end{aligned}$$

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

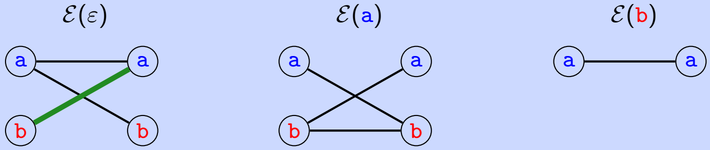


Extension graphs

The *extension graph* of a word $w \in \mathcal{L}$ is the undirected bipartite graph $\mathcal{E}(w)$ with vertices $L(w) \sqcup R(w)$ and edges $B(w)$, where

$$\begin{array}{rcl} L(w) & = & \{u \in \mathcal{A} \mid uw \in \mathcal{L}\} \\ R(w) & = & \{v \in \mathcal{A} \mid wv \in \mathcal{L}\} \\ B(w) & = & \{(u,v) \in \mathcal{A} \times \mathcal{A} \mid uwv \in \mathcal{L}\} \end{array}$$

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



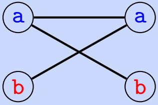
Extension graphs

The *extension graph* of a word $w \in \mathcal{L}$ is the undirected bipartite graph $\mathcal{E}(w)$ with vertices $L(w) \sqcup R(w)$ and edges $B(w)$, where

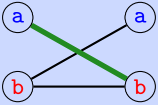
$$\begin{aligned} L(w) &= \{u \in \mathcal{A} \mid uw \in \mathcal{L}\} \\ R(w) &= \{v \in \mathcal{A} \mid wv \in \mathcal{L}\} \\ B(w) &= \{(u, v) \in \mathcal{A} \times \mathcal{A} \mid uwv \in \mathcal{L}\} \end{aligned}$$

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

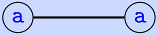
$\mathcal{E}(\varepsilon)$



$\mathcal{E}(\mathrm{a})$



$\mathcal{E}(\mathrm{b})$

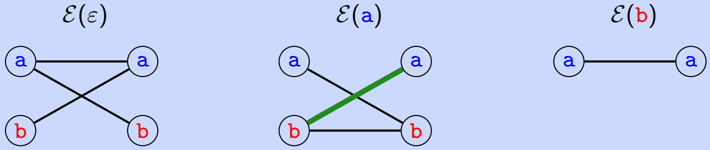


Extension graphs

The *extension graph* of a word $w \in \mathcal{L}$ is the undirected bipartite graph $\mathcal{E}(w)$ with vertices $L(w) \sqcup R(w)$ and edges $B(w)$, where

$$\begin{aligned} L(w) &= \{u \in \mathcal{A} \mid uw \in \mathcal{L}\} \\ R(w) &= \{v \in \mathcal{A} \mid wv \in \mathcal{L}\} \\ B(w) &= \{(u,v) \in \mathcal{A} \times \mathcal{A} \mid uwv \in \mathcal{L}\} \end{aligned}$$

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



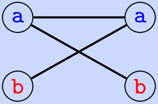
Extension graphs

The *extension graph* of a word $w \in \mathcal{L}$ is the undirected bipartite graph $\mathcal{E}(w)$ with vertices $L(w) \sqcup R(w)$ and edges $B(w)$, where

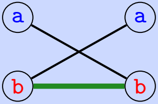
$$\begin{aligned} L(w) &= \{u \in \mathcal{A} \mid uw \in \mathcal{L}\} \\ R(w) &= \{v \in \mathcal{A} \mid wv \in \mathcal{L}\} \\ B(w) &= \{(u,v) \in \mathcal{A} \times \mathcal{A} \mid uwv \in \mathcal{L}\} \end{aligned}$$

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

$\mathcal{E}(\varepsilon)$



$\mathcal{E}(\mathrm{a})$



$\mathcal{E}(\mathrm{b})$

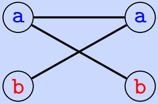
Extension graphs

The *extension graph* of a word $w \in \mathcal{L}$ is the undirected bipartite graph $\mathcal{E}(w)$ with vertices $L(w) \sqcup R(w)$ and edges $B(w)$, where

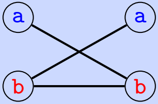
$$\begin{aligned} L(w) &= \{u \in \mathcal{A} \mid uw \in \mathcal{L}\} \\ R(w) &= \{v \in \mathcal{A} \mid wv \in \mathcal{L}\} \\ B(w) &= \{(u, v) \in \mathcal{A} \times \mathcal{A} \mid uwv \in \mathcal{L}\} \end{aligned}$$

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

$\mathcal{E}(\varepsilon)$



$\mathcal{E}(\mathtt{a})$



$\mathcal{E}(\mathtt{b})$

Extension graphs

The *extension graph* of a word $w \in \mathcal{L}$ is the undirected bipartite graph $\mathcal{E}(w)$ with vertices $L(w) \sqcup R(w)$ and edges $B(w)$, where

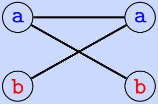
$$\begin{aligned} L(w) &= \{u \in \mathcal{A} \mid uw \in \mathcal{L}\} \\ R(w) &= \{v \in \mathcal{A} \mid wv \in \mathcal{L}\} \\ B(w) &= \{(u, v) \in \mathcal{A} \times \mathcal{A} \mid uwv \in \mathcal{L}\} \end{aligned}$$

The *multiplicity* of a word w is the quantity

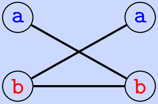
$$m(w) = \text{Card}\,(B(w)) - \text{Card}\,(L(w)) - \text{Card}\,(R(w)) + 1.$$

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

$\mathcal{E}(\varepsilon)$



$\mathcal{E}(\mathtt{a})$



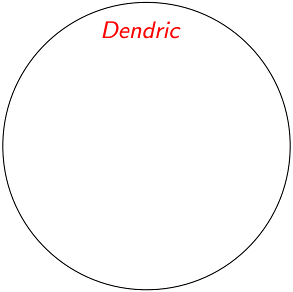
$\mathcal{E}(\mathtt{b})$

$m(\mathtt{a}) = 3 - 2 - 2 + 1 = 0$

Dendric and neutral languages

Definition

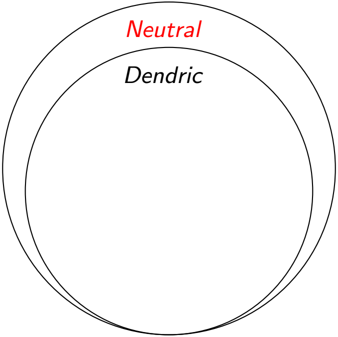
A language $\mathcal{L}$ is called (purely) *dendric* if the graph $\mathcal{E}(w)$ is a tree for any $w \in \mathcal{L}$.



Dendric and neutral languages

Definition

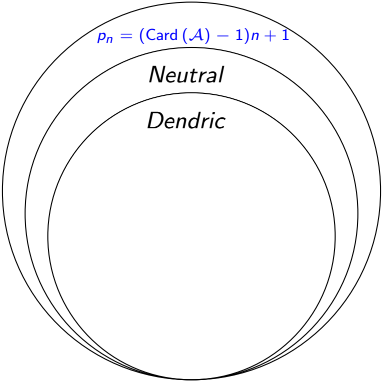
A language $\mathcal{L}$ is called (purely) *dendric* if the graph $\mathcal{E}(w)$ is a tree for any $w \in \mathcal{L}$. It is called *neutral* if every word w has multiplicity $m(w) = 0$.



Dendric and neutral languages

Definition

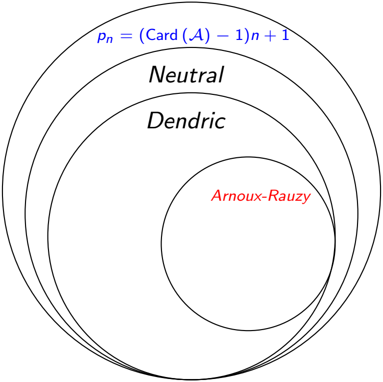
A language $\mathcal{L}$ is called (purely) *dendric* if the graph $\mathcal{E}(w)$ is a tree for any $w \in \mathcal{L}$. It is called *neutral* if every word w has multiplicity $m(w) = 0$.



Dendric and neutral languages

Definition

A language $\mathcal{L}$ is called (purely) *dendric* if the graph $\mathcal{E}(w)$ is a tree for any $w \in \mathcal{L}$. It is called *neutral* if every word w has multiplicity $m(w) = 0$.

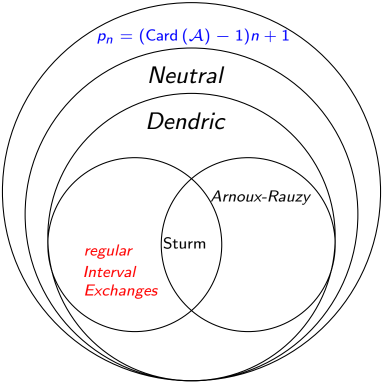


[Berthé, De Felice, D., Leroy, Perrin, Reutenauer, Rindone: "**Acyclic, connected and tree sets**" (2014)]

Dendric and neutral languages

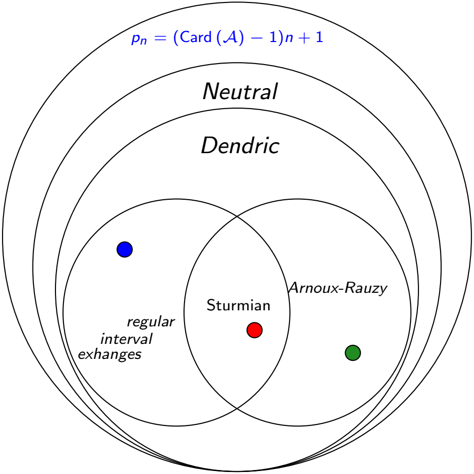
Definition

A language $\mathcal{L}$ is called (purely) *dendric* if the graph $\mathcal{E}(w)$ is a tree for any $w \in \mathcal{L}$. It is called *neutral* if every word w has multiplicity $m(w) = 0$.



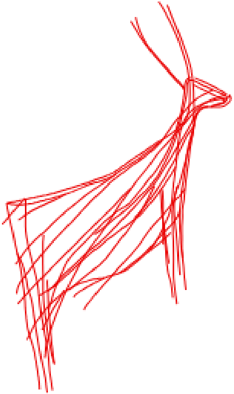
[Berthé, De Felice, D., Leroy, Perrin, Reutenauer, Rindone: "**Bifix codes and interval exchanges**" (2015)]

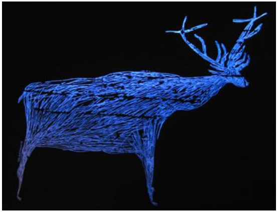
Dendric and neutral languages



- • Fibonacci
- ? 2-coded Fibonacci
- • Tribonacci
- ? 2-coded Tribonacci
- • regular IE
- ? 2-coded regular IE

Bifix codes





Bifix codes

Definition

A *bifix code* is a set $B \subset \mathcal{A}^+$ of nonempty words that does not contain any proper prefix or suffix of its elements.

Example

- ✓ $\{aa, ab, ba\}$
- ✓ $\{aa, ab, bba, bbb\}$
- ✓ $\{ac, bcc, bcbca\}$
- ✗ { por, portugal, vinho }
- ✗ { saudade, do, fado }
- ✗ { do, douro, ouro }

Bifix codes

Definition

A *bifix code* is a set $B \subset \mathcal{A}^+$ of nonempty words that does not contain any proper prefix or suffix of its elements.

A bifix code $B \subset \mathcal{L}$ is *$\mathcal{L}$-maximal* if it is not properly contained in a bifix code $C \subset \mathcal{L}$.

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

The set $B = \{\mathtt{aa}, \mathtt{ab}, \mathtt{ba}\}$ is an $\mathcal{L}$-maximal bifix code.
It is not an $\mathcal{A}^*$-maximal bifix code, since $B \subset B \cup \{\mathtt{bb}\}$.

Bifix codes

Definition

A *bifix code* is a set $B \subset \mathcal{A}^+$ of nonempty words that does not contain any proper prefix or suffix of its elements.

A bifix code $B \subset \mathcal{L}$ is *$\mathcal{L}$-maximal* if it is not properly contained in a bifix code $C \subset \mathcal{L}$.

A *coding morphism* for a bifix code $B \subset A^+$ is a morphism $f : \mathcal{B}^* \to \mathcal{A}^*$ which maps bijectively $\mathcal{B}$ onto B.

Example

The map $f : \{\mathtt{u}, \mathtt{v}, \mathtt{w}\}^* \to \{\mathtt{a}, \mathtt{b}\}^*$ is a coding morphism for $B = \{\mathtt{aa}, \mathtt{ab}, \mathtt{ba}\}$.

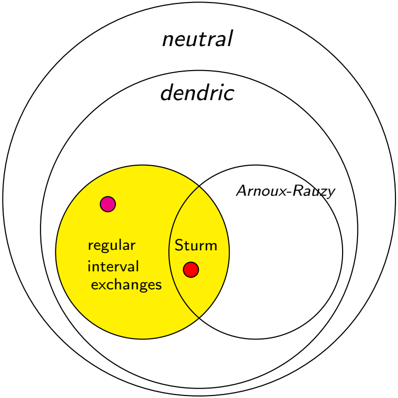
$$f : \begin{cases} \mathtt{u} \mapsto \mathtt{aa} \\ \mathtt{v} \mapsto \mathtt{ab} \\ \mathtt{w} \mapsto \mathtt{ba} \end{cases}$$

When $\mathcal{L}$ is factorial and B is an $\mathcal{L}$-maximal bifix code, the set $f^{-1}(\mathcal{L})$ is called a *maximal bifix decoding* of $\mathcal{L}$.

Maximal bifix decoding

Theorem

The family of regular interval exchanges languages is closed under maximal bifix decoding.



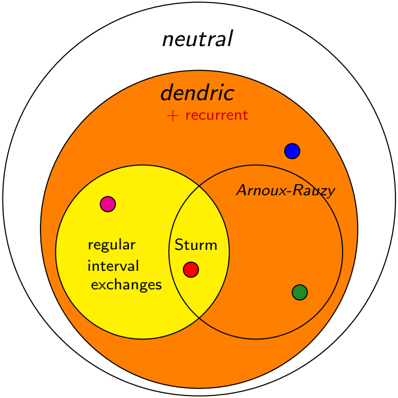
- Fibonacci
- 2-coded Fibonacci

[Berthé, De Felice, D., Leroy, Perrin, Reutenauer, Rindone: "**Bifix codes and interval exchanges**" (2015)]

Maximal bifix decoding

Theorem

The family of *recurrent* dendric languages is closed under maximal bifix decoding.



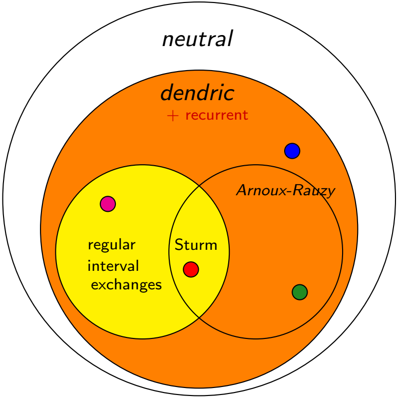
- Fibonacci
- 2-coded Fibonacci
- Tribonacci
- 2-coded Tribonacci

[Berthé, De Felice, D., Leroy, Perrin, Reutenauer, Rindone: "**Maximal bifix decoding**" (2015)]

Maximal bifix decoding

Theorem

The family of *recurrent* dendric languages is closed under maximal bifix decoding.



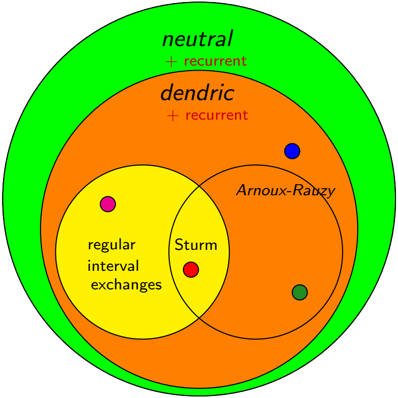
- Fibonacci
- 2-coded Fibonacci
- Tribonacci
- 2-coded Tribonacci

Theorem. All complete bifix decodings of uniformly recurrent dendric languages are also uniformly recurrent [Costa (2023)]

Maximal bifix decoding

Theorem

The family of *recurrent* neutral languages is closed under maximal bifix decoding.



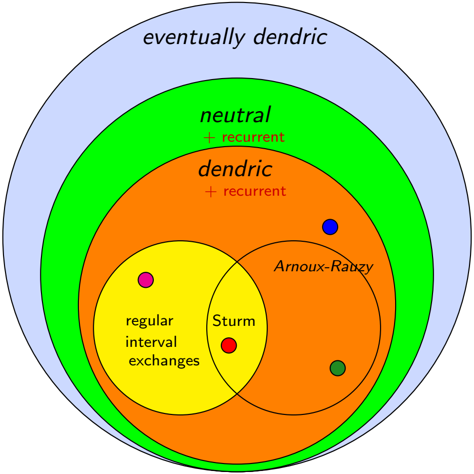
- Fibonacci
- 2-coded Fibonacci
- Tribonacci
- 2-coded Tribonacci

[D., Perrin "**Neutral and tree sets of arbitrary characteristic**" (2016)]

Maximal bifix decoding

Theorem

The family of **eventually** dendric languages is closed under maximal bifix decoding.



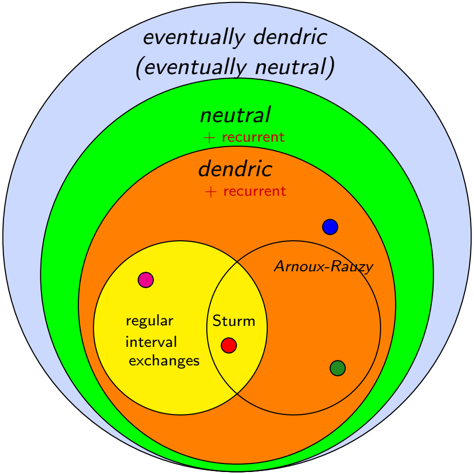
- Fibonacci
- 2-coded Fibonacci
- Tribonacci
- 2-coded Tribonacci

[D., Perrin "**Eventually dendric shift spaces**" (2019)]

Maximal bifix decoding

Theorem

The family of **eventually** dendric languages is closed under maximal bifix decoding.



- Fibonacci
- 2-coded Fibonacci
- Tribonacci
- 2-coded Tribonacci

[Gheeraert "**Some properties of morphic images of (eventually) dendric words**" (2023)]

Parse and degree

Definition

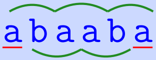
A *parse* of a word w with respect to a bifix code B is a triple (p, x, s) such that:

- $w = pxs$,
- p has no suffix in B,
- $x \in X^*$ and
- s has no prefix in B.

Example

Let $B = \{\mathtt{aa}, \mathtt{ab}, \mathtt{ba}\}$ and $w = \mathtt{abaaba}$. The two possible parses of w are:

- $(\varepsilon, \mathtt{ab\,aa\,ba}, \varepsilon)$,
- $(\mathtt{a}, \mathtt{ba\,ab}, \mathtt{a})$.



Parse and degree

Definition

A *parse* of a word w with respect to a bifix code B is a triple (p, x, s) such that:

- $w = pxs$,
- p has no suffix in B,
- $x \in X^*$ and
- s has no prefix in B.

The $\mathcal{L}$-*degree* of B is the maximal number of parses with respect to B of a word in $\mathcal{L}$.

Example (Fibonacci)

- The set $B = \{\mathtt{aa}, \mathtt{ab}, \mathtt{ba}\}$ has $\mathcal{L}$-degree 2.
- The set $\mathcal{L} \cap \mathcal{A}^n$ has $\mathcal{L}$-degree n.

Cardinality of bifix codes

Theorem

Let $\mathcal{L}$ be a recurrent neutral set.
For any finite $\mathcal{L}$-maximal bifix code B of $\mathcal{L}$-degree n, one has

$$\mathrm{Card}\,(B) = n\big(\mathrm{Card}\,(\mathcal{A}) - 1\big) + 1.$$

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

The three possible $\mathcal{L}$-maximal bifix codes of $\mathcal{L}$-degree 2 are :

- $\{\mathrm{aa}, \mathrm{ab}, \mathrm{ba}\}$
- $\{\mathrm{a}, \mathrm{baab}, \mathrm{bab}\}$
- $\{\mathrm{aa}, \mathrm{aba}, \mathrm{b}\}$

Each of them has cardinality $3 = 2(2-1) + 1$.

Cardinality of bifix codes

Theorem

Let $\mathcal{L}$ be a recurrent neutral set.
For any finite $\mathcal{L}$-maximal bifix code B of $\mathcal{L}$-degree n, one has

$$\mathrm{Card}\,(B) = n\big(\mathrm{Card}\,(\mathcal{A}) - 1\big) + 1.$$

Theorem

Let $\mathcal{L}$ be a *uniformly* recurrent set.
If every finite $\mathcal{L}$-maximal bifix code of $\mathcal{L}$-degree n has $n\big(\mathrm{Card}\,(A) - 1\big) + 1$ elements, then $\mathcal{L}$ is neutral.

Finite index basis property

Example (Fibonacci)

The $\mathcal{L}$-maximal bifix code $B = \{\mathtt{aa}, \mathtt{ab}, \mathtt{ba}\}$ of $\mathcal{L}$-degree 2 is a basis of $\langle \mathcal{A}^2 \rangle$. Indeed

$$\mathtt{bb} = \mathtt{ba}\ (\mathtt{aa})^{-1}\ \mathtt{ab}$$

Finite index basis property

Example (Fibonacci)

The $\mathcal{L}$-maximal bifix code $B = \{\mathtt{aa}, \mathtt{ab}, \mathtt{ba}\}$ of $\mathcal{L}$-degree 2 is a basis of $\langle \mathcal{A}^2 \rangle$. Indeed

$$\mathtt{bb} = \mathtt{ba}\ (\mathtt{aa})^{-1}\ \mathtt{ab}$$

Also $\mathcal{L} \cap \mathcal{A}^3 = \{\mathtt{aab}, \mathtt{aba}, \mathtt{baa}, \mathtt{bab}\}$ is a basis of $\langle \mathcal{A}^3 \rangle$:

$$\begin{aligned} \mathtt{aaa} &= \mathtt{aab}\,(\mathtt{bab})^{-1}\,\mathtt{baa} \\ \mathtt{abb} &= \mathtt{aba}\,(\mathtt{baa})^{-1}\,\mathtt{bab} \\ \mathtt{bba} &= \mathtt{bab}\,(\mathtt{aab})^{-1}\,\mathtt{aba} \\ \mathtt{bbb} &= \mathtt{bba}\,(\mathtt{aba})^{-1}\,\mathtt{aab} \end{aligned}$$

Finite index basis property

Example (Fibonacci)

The $\mathcal{L}$-maximal bifix code $B = \{\mathtt{aa}, \mathtt{ab}, \mathtt{ba}\}$ of $\mathcal{L}$-degree 2 is a basis of $\langle \mathcal{A}^2 \rangle$. Indeed

$$\mathtt{bb} = \mathtt{ba}\ (\mathtt{aa})^{-1}\ \mathtt{ab}$$

Also $\mathcal{L} \cap \mathcal{A}^3 = \{\mathtt{aab}, \mathtt{aba}, \mathtt{baa}, \mathtt{bab}\}$ is a basis of $\langle \mathcal{A}^3 \rangle$:

$$\begin{aligned} \mathtt{aaa} &= \mathtt{aab}\,(\mathtt{bab})^{-1}\,\mathtt{baa} \\ \mathtt{abb} &= \mathtt{aba}\,(\mathtt{baa})^{-1}\,\mathtt{bab} \\ \mathtt{bba} &= \mathtt{bab}\,(\mathtt{aab})^{-1}\,\mathtt{aba} \\ \mathtt{bbb} &= \mathtt{bba}\,(\mathtt{aba})^{-1}\,\mathtt{aab} \end{aligned}$$

$\{\mathtt{aa}, \mathtt{aba}, \mathtt{b}\}$ of $\mathcal{L}$-degree 2 and $[\langle \mathtt{aa}, \mathtt{aba}, \mathtt{b} \rangle : \mathbb{F}_{\mathcal{A}}] = 2$

Finite index basis property

Definition

A set $\mathcal{L} \subset \mathcal{A}^+$ satisfies the *finite index basis property* if for any finite bifix code $B \subset \mathcal{L}$:

B is an $\mathcal{L}$-maximal bifix code of $\mathcal{L}$-degree d if and only if it is a basis of a subgroup of index d of the free group on $\mathbb{F}_{\mathcal{A}}$.

Finite index basis property

Definition

A set $\mathcal{L} \subset \mathcal{A}^+$ satisfies the *finite index basis property* if for any finite bifix code $B \subset \mathcal{L}$:

B is an $\mathcal{L}$-maximal bifix code of $\mathcal{L}$-degree d if and only if it is a basis of a subgroup of index d of the free group on $\mathbb{F}_{\mathcal{A}}$.

Theorem

An Arnoux-Rauzy set satisfies the finite index basis property.

Finite index basis property

Definition

A set $\mathcal{L} \subset \mathcal{A}^+$ satisfies the *finite index basis property* if for any finite bifix code $B \subset \mathcal{L}$:

B is an $\mathcal{L}$-maximal bifix code of $\mathcal{L}$-degree d if and only if it is a basis of a subgroup of index d of the free group on $\mathbb{F}_{\mathcal{A}}$.

Theorem

A regular interval exchange set satisfies the finite index basis property.

Finite index basis property

Definition

A set $\mathcal{L} \subset \mathcal{A}^+$ satisfies the *finite index basis property* if for any finite bifix code $B \subset \mathcal{L}$:

B is an $\mathcal{L}$-maximal bifix code of $\mathcal{L}$-degree d if and only if it is a basis of a subgroup of index d of the free group on $\mathbb{F}_{\mathcal{A}}$.

Theorem

A (*uniformly*) *recurrent* dendric set satisfies the finite index basis property.

Finite index basis property

Definition

A set $\mathcal{L} \subset \mathcal{A}^+$ satisfies the *finite index basis property* if for any finite bifix code $B \subset \mathcal{L}$:

B is an $\mathcal{L}$-maximal bifix code of $\mathcal{L}$-degree d if and only if it is a basis of a subgroup of index d of the free group on $\mathbb{F}_{\mathcal{A}}$.

Theorem

A (*uniformly*) *recurrent* dendric set satisfies the finite index basis property.

Theorem

A *uniformly recurrent* set satisfying the finite index basis property is a dendric set.

Return words

A (*right*) *return word* to w in $\mathcal{L}$ is a nonempty word u such that $wu \in \mathcal{L}$ starts and ends with w but has no w as an internal factor. Formally,

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

Example (Fibonacci)

$\mathcal{R}(\mathtt{b}) = \{\mathtt{a\underline{b}}, \mathtt{aa\underline{b}}\}$

$$\varphi(\mathtt{a})^\omega = \mathtt{abaababaabaababaababaabaababaabaab}\cdots$$

$\mathcal{R}(\mathtt{aa}) = \{\mathtt{b\underline{aa}}, \mathtt{bab\underline{aa}}\}$

$$\varphi(\mathtt{a})^\omega = \mathtt{abaababaabaababaababaabaababaabaab}\cdots$$

The Return Theorem

Theorem

Let $\mathcal{L}$ be a recurrent dendric language and $w \in \mathcal{L}$.
Then $\mathcal{R}(w)$ is a basis of the free group $\mathbb{F}_{\mathcal{A}}$.

Example (Fibonacci)

The set $\mathcal{R}(\mathtt{b}) = \{\mathtt{ab}, \mathtt{aab}\}$ is a basis of the free group. Indeed,

$$\begin{aligned} \mathtt{a} &= \mathtt{aab}\,(\mathtt{ab})^{-1} \\ \mathtt{b} &= \mathtt{a}^{-1}\,\mathtt{ab} \end{aligned}$$

$$\langle \mathcal{R}(\mathtt{b}) \rangle = \langle \mathtt{ab}, \mathtt{aab} \rangle = \langle \mathtt{a}, \mathtt{b} \rangle = \mathbb{F}_{\mathcal{A}}$$

The Return Theorem

Theorem

Let $\mathcal{L}$ be a recurrent dendric language and $w \in \mathcal{L}$.
Then $\mathcal{R}(w)$ is a basis of the free group $\mathbb{F}_{\mathcal{A}}$.

Example (Fibonacci)

The set $\mathcal{R}(\mathtt{aa}) = \{\mathtt{aab}, \mathtt{aabab}\}$ is a basis of the free group. Indeed,

$$\begin{aligned} \mathtt{a} &= \mathtt{aab}\,(\mathtt{aabab})^{-1}\,\mathtt{aab} \\ \mathtt{b} &= \mathtt{a}^{-1}\,\mathtt{a}^{-1}\,\mathtt{aab} \end{aligned}$$

$$\langle \mathcal{R}(\mathtt{aa}) \rangle = \langle \mathtt{aab}, \mathtt{aabab} \rangle = \langle \mathtt{a}, \mathtt{b} \rangle = \mathbb{F}_{\mathcal{A}}$$

The Return Theorem

Theorem

Let $\mathcal{L}$ be a recurrent dendric language and $w \in \mathcal{L}$.
Then $\mathcal{R}(w)$ is a basis of the free group $\mathbb{F}_{\mathcal{A}}$.

Corollary

For every $w \in \mathcal{L}$, we have $\mathrm{Card}\,(\mathcal{R}(w)) = \mathrm{Card}\,(\mathcal{A})$

The Return Theorem

Theorem

Let $\mathcal{L}$ be a recurrent dendric language and $w \in \mathcal{L}$.
Then $\mathcal{R}(w)$ is a basis of the free group $\mathbb{F}_{\mathcal{A}}$.

Corollary

For every $w \in \mathcal{L}$, we have $\text{Card}\,(\mathcal{R}(w)) = \text{Card}\,(\mathcal{A})$

Theorem

Let $\mathcal{L}$ be a recurrent connected language.
For any $w \in \mathcal{L}$, the set $\mathcal{R}(w)$ generates the free group $\mathbb{F}_{\mathcal{A}}$.

The Return Theorem

Theorem

Let $\mathcal{L}$ be a recurrent dendric language and $w \in \mathcal{L}$.
Then $\mathcal{R}(w)$ is a basis of the free group $\mathbb{F}_{\mathcal{A}}$.

Corollary

For every $w \in \mathcal{L}$, we have $\text{Card}\,(\mathcal{R}(w)) = \text{Card}\,(\mathcal{A})$

Theorem

Let $\mathcal{L}$ be a recurrent suffix-connected language.
For any $w \in \mathcal{L}$, the set $\mathcal{R}(w)$ generates the free group $\mathbb{F}_{\mathcal{A}}$.

[Goulet-Ouellet: "**Suffix-connected languages**" (2022)]

Open problems

- Is the class of suffix-connected languages closed under complete bifix decoding?

- Is it possible to characterize the languages such that every set of return words generates (resp., is a basis of) the free group?

- Are dendric languages rigid?

  An infinite word $\mathbf{x}$ is *rigid* if $\mathrm{Stab}(\mathbf{x}) = \{\sigma : \mathcal{A}^* \to \mathcal{A}^* \mid \sigma(\mathbf{x}) = \mathbf{x}\}$ is cyclic.

- Connection with profinite algebra

  Almeida, Costa (2016), Almeida, Costa, Kyriakoglou, Perrin (2020), Goulet-Ouellet (2022), Costa (2023)

Obrigado pela sua atenção