Breadth-first signature and numeration systems

Victor Marsault,
joint work with Jacques Sakarovitch

CNRS / Telecom-ParisTech, Paris, France

Séminaire Automate, Paris,
2014–10–17

Motivations

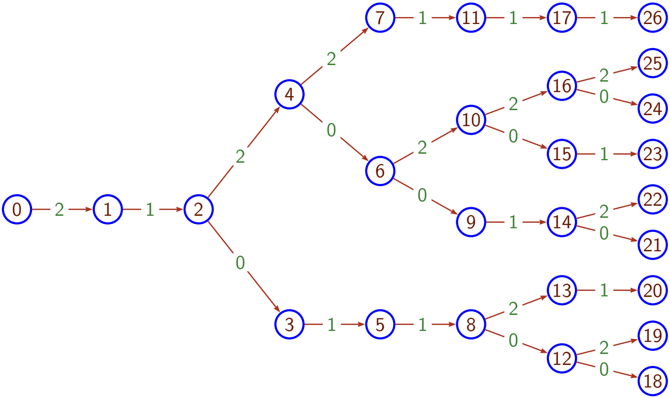


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Motivations

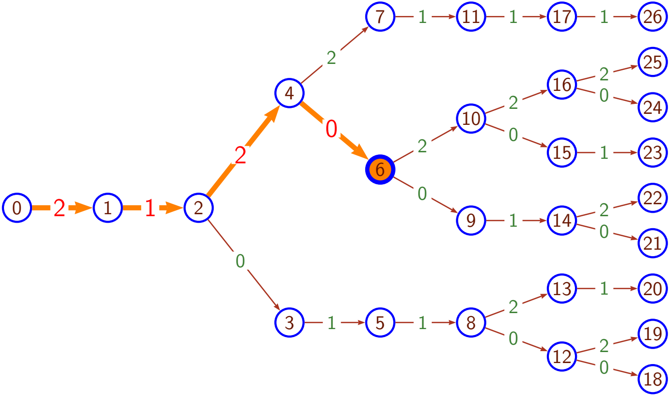


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Motivations

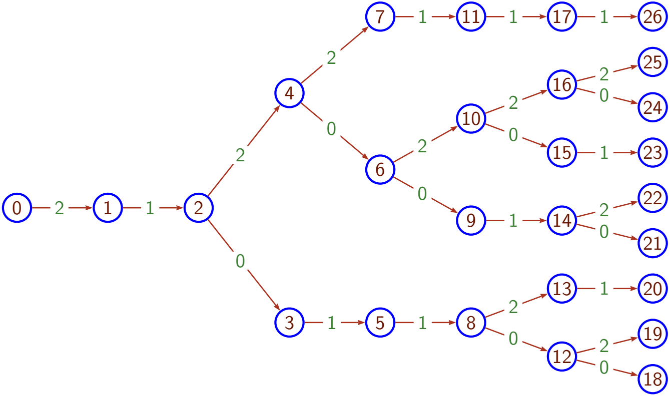


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

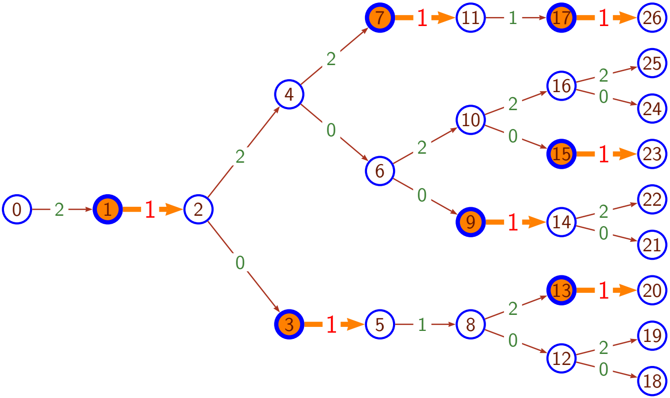


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

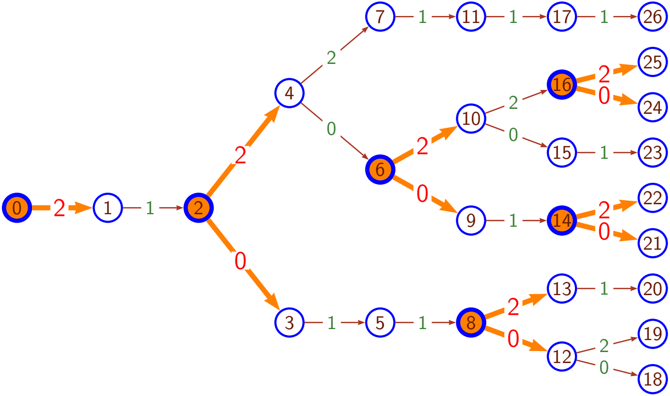


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

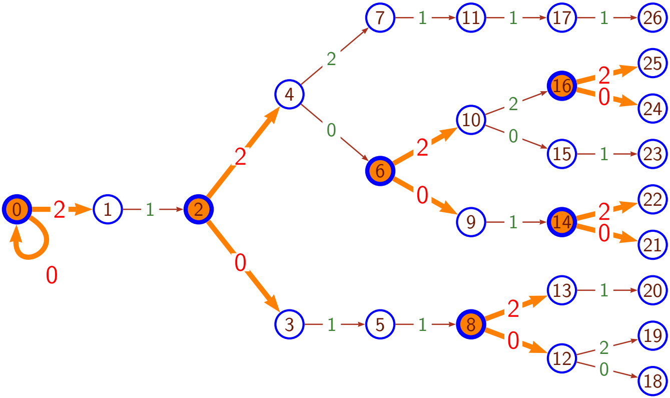


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

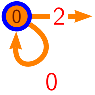


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

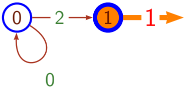


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

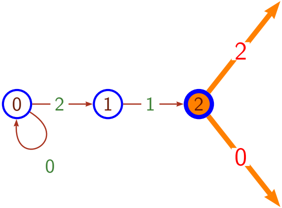


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

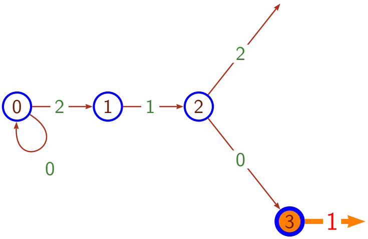


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

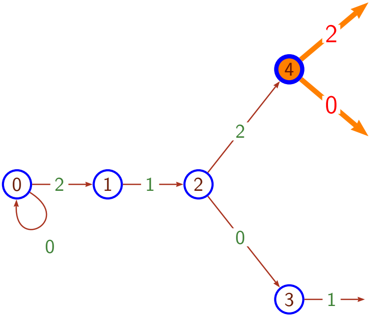


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

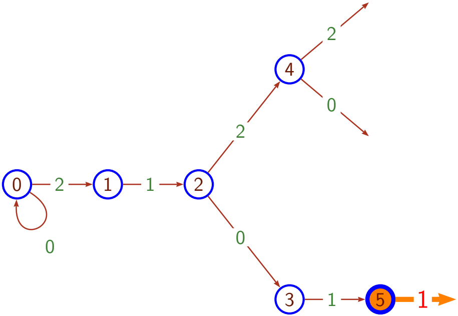


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

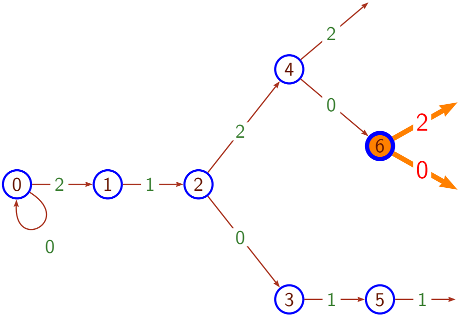


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

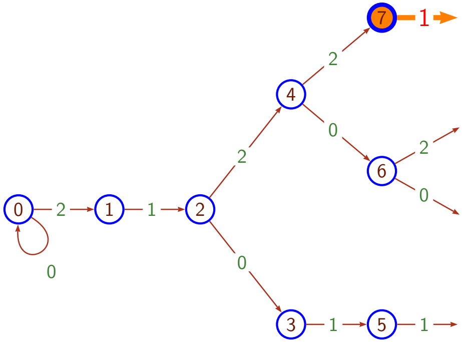


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

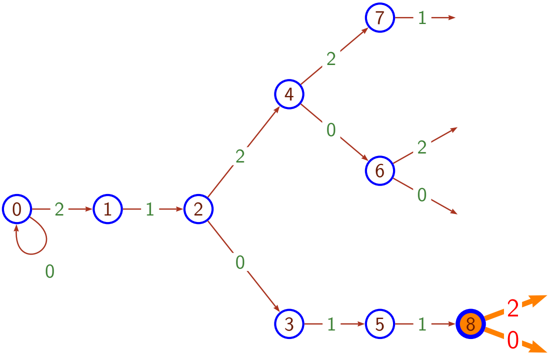


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

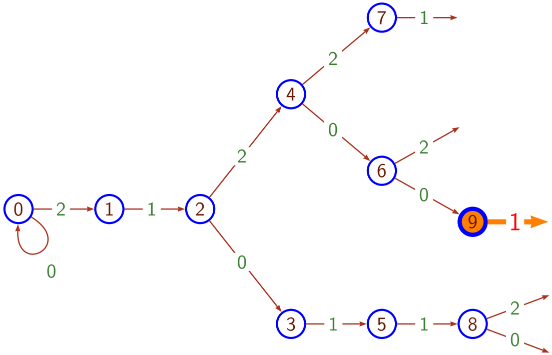


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

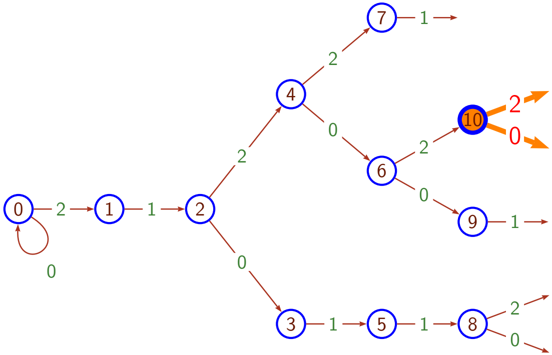


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

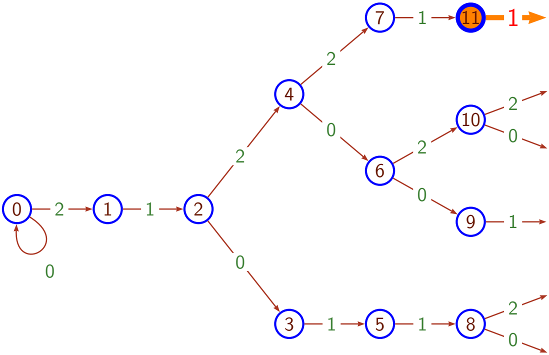


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

Motivations

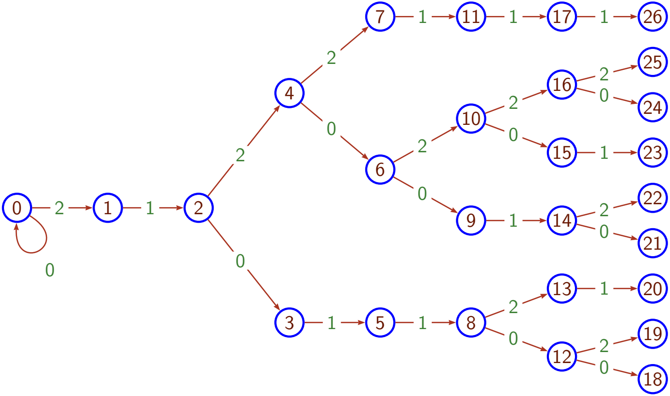


FIGURE: $L_{\frac{3}{2}}$, the language of integer representations in base $\frac{3}{2}$.

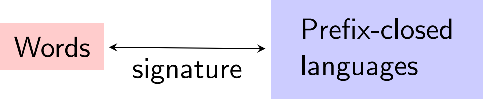
Proposition (Akiyama Frougny Sakarovitch 2008)

$L_{\frac{p}{q}}$ is not context-free.

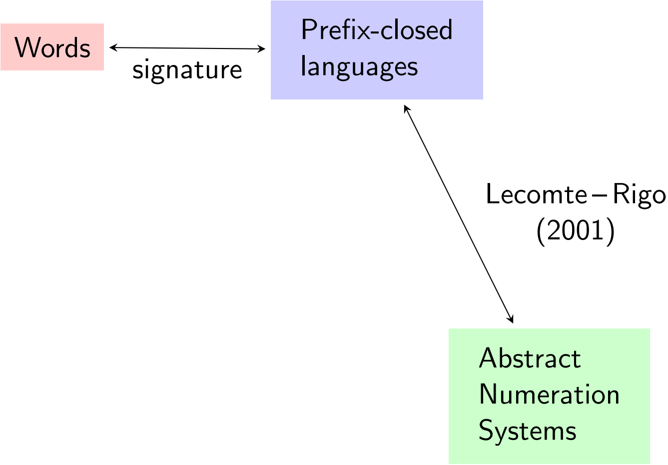
Overview

Prefix-closed languages

Overview



Overview



Words

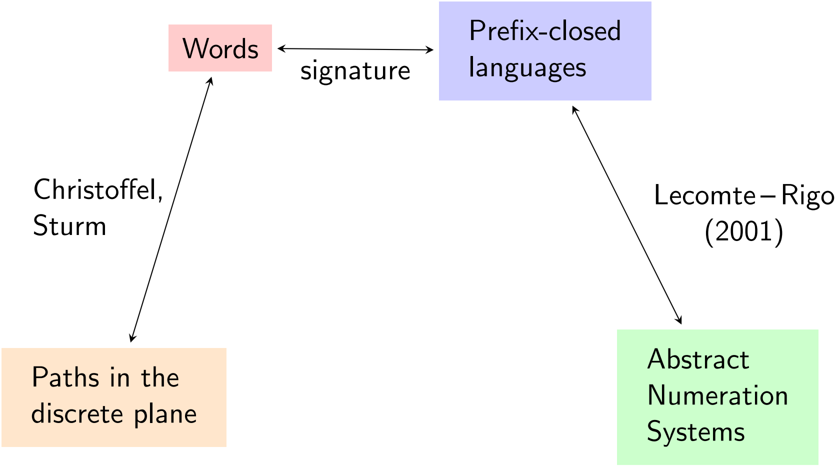
signature

Prefix-closed languages

Lecomte – Rigo (2001)

Abstract Numeration Systems

Overview



Words

signature

Prefix-closed languages

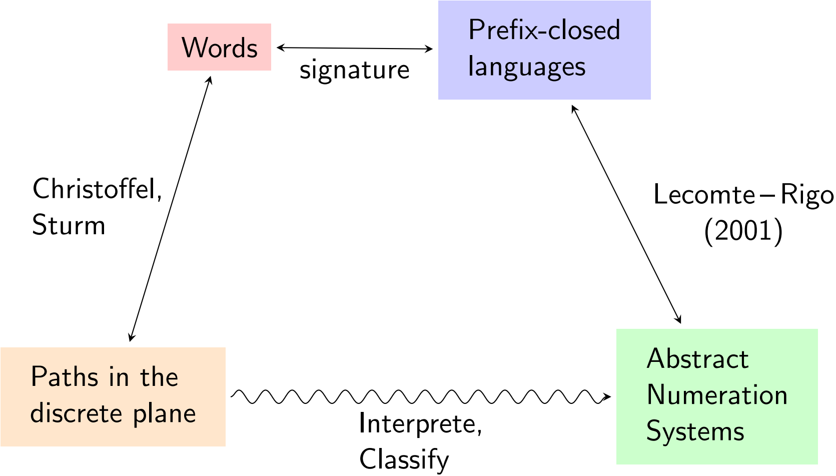
Christoffel, Sturm

Lecomte – Rigo (2001)

Paths in the discrete plane

Abstract Numeration Systems

Overview



Words

Prefix-closed languages

signature

Christoffel, Sturm

Lecomte – Rigo (2001)

Paths in the discrete plane

Abstract Numeration Systems

Interprete, Classify

Outline

We call *tree* a...

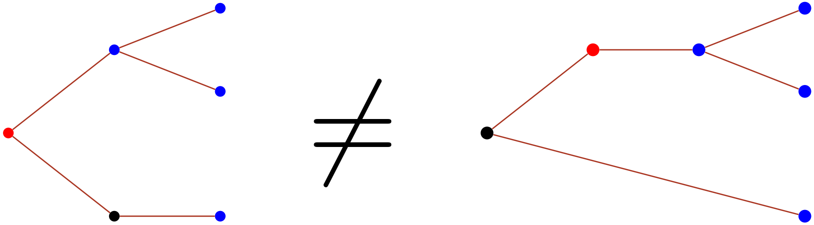
Directed graph which is

- **Rooted:** a node is called *the root* (leftmost in the figures)
- **Directed outward from the root:** there is a unique path from the root to every other node.
- **Ordered:** the children of every node are ordered
  (In the figures, lower children are smaller.)

We call *tree* a...

Directed graph which is

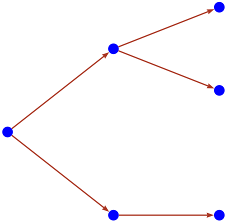
- **Rooted:** a node is called *the root* (leftmost in the figures)
- **Directed outward from the root:** there is a unique path from the root to every other node.
- **Ordered:** the children of every node are ordered
  (In the figures, lower children are smaller.)



We call *tree* a...

Directed graph which is

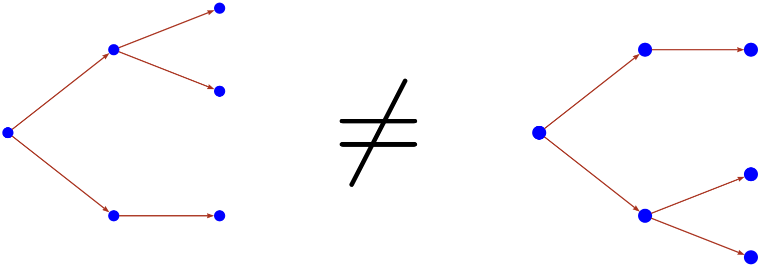
- **Rooted:** a node is called *the root* (leftmost in the figures)
- **Directed outward from the root:** there is a unique path from the root to every other node.
- **Ordered:** the children of every node are ordered (In the figures, lower children are smaller.)



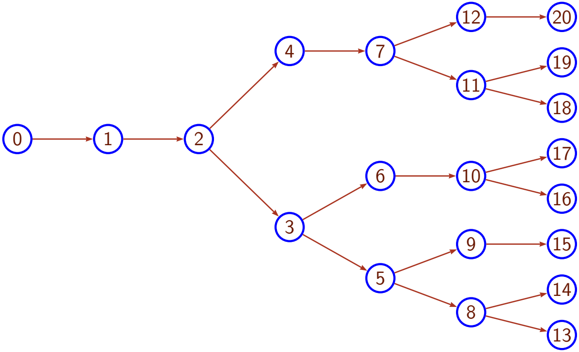
We call *tree* a...

Directed graph which is

- **Rooted:** a node is called *the root* (leftmost in the figures)
- **Directed outward from the root:** there is a unique path from the root to every other node.
- **Ordered:** the children of every node are ordered
  (In the figures, lower children are smaller.)

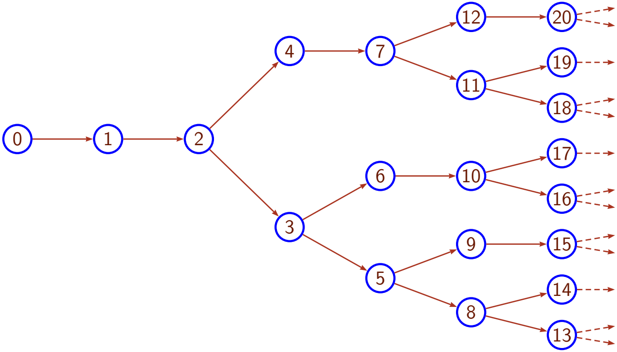


Each tree has a canonical breadth-first traversal



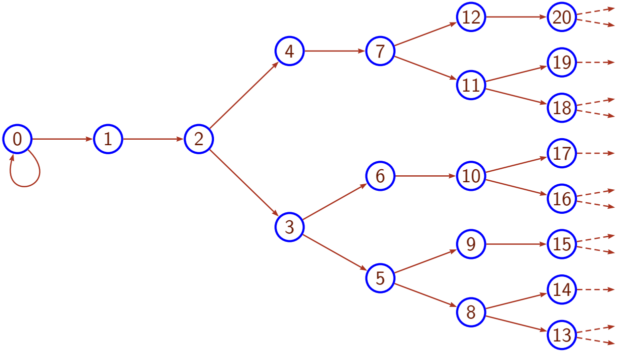
Two more features

- We consider infinite trees only.



Two more features

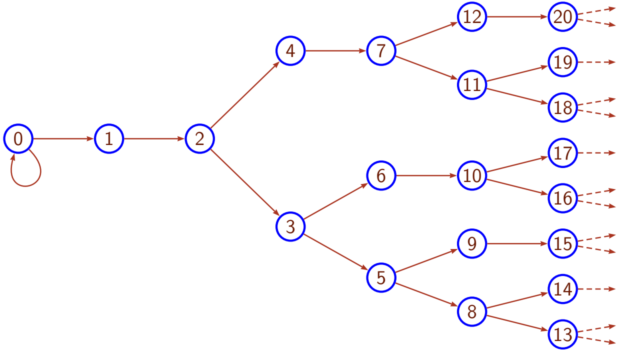
- We consider infinite trees only.
- For convenience, there is loop on the root.



Signature of a tree

Definition

The **signature** of a tree is the sequence of the degree of the nodes taken in breadth-first order.

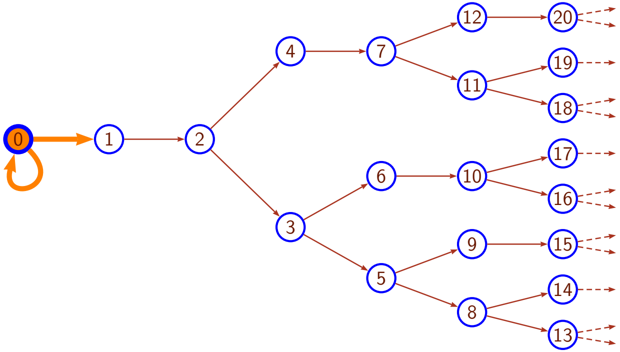


s =

Signature of a tree

Definition

The **signature** of a tree is the sequence of the degree of the nodes taken in breadth-first order.

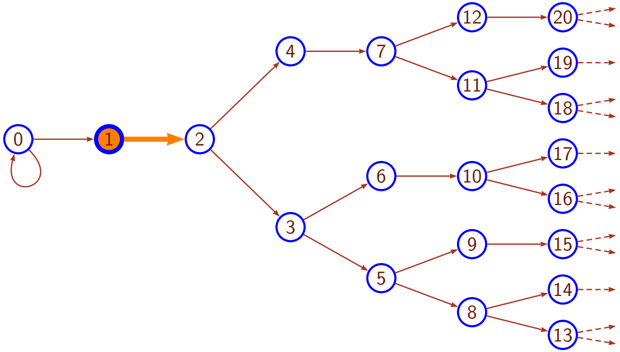


$\mathbf{s} = \boxed{2}$

Signature of a tree

Definition

The **signature** of a tree is the sequence of the degree of the nodes taken in breadth-first order.

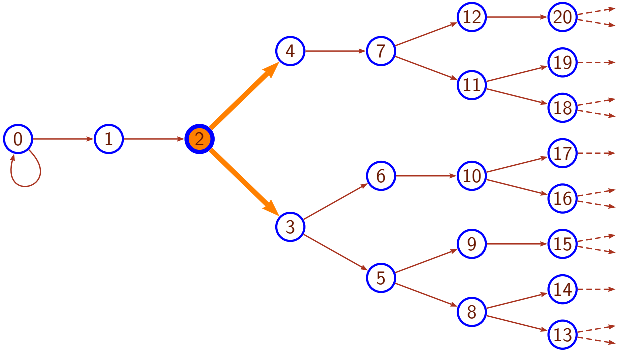


$$\mathbf{s} = 2\ 1$$

Signature of a tree

Definition

The **signature** of a tree is the sequence of the degree of the nodes taken in breadth-first order.

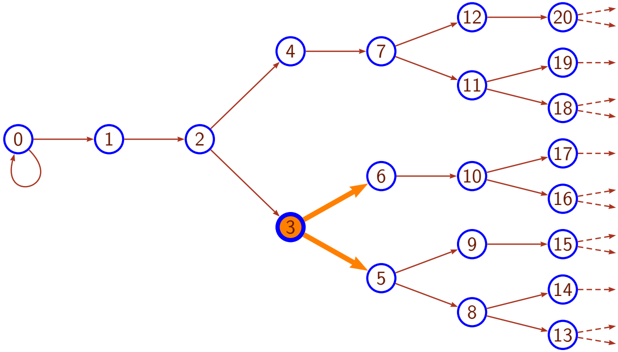


$$\mathbf{s} = 2\ 1\ 2$$

Signature of a tree

Definition

The **signature** of a tree is the sequence of the degree of the nodes taken in breadth-first order.

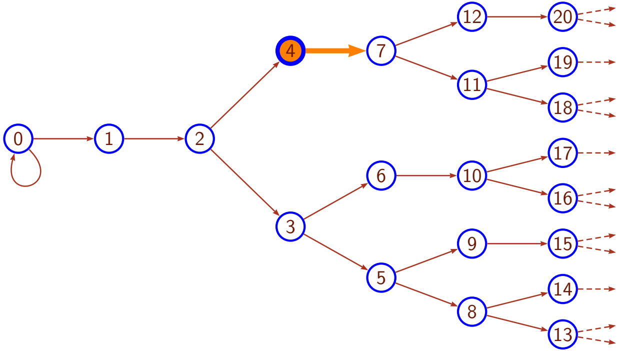


$\mathbf{s} = 2\ 1\ 2\ 2$

Signature of a tree

Definition

The **signature** of a tree is the sequence of the degree of the nodes taken in breadth-first order.

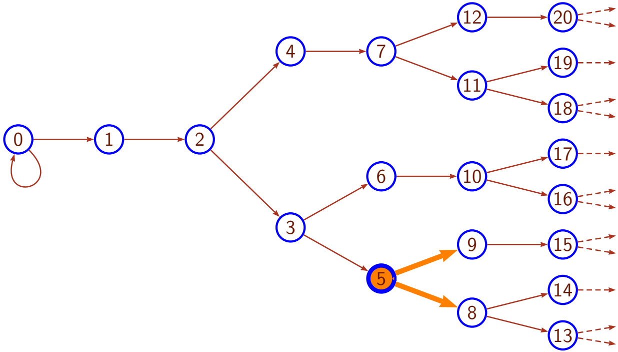


$$\mathbf{s} = 2\ 1\ 2\ 2\ 1$$

Signature of a tree

Definition

The **signature** of a tree is the sequence of the degree of the nodes taken in breadth-first order.

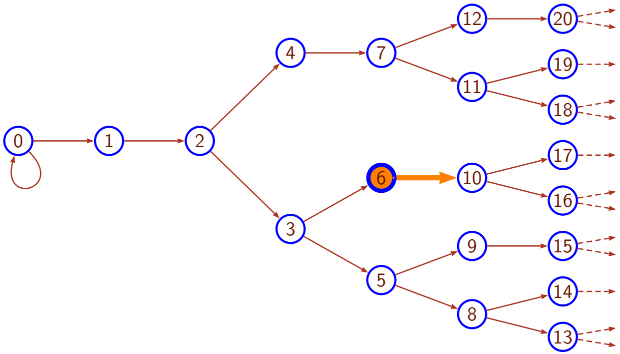


$$\mathbf{s} = 2\ 1\ 2\ 2\ 1\ 2$$

Signature of a tree

Definition

The **signature** of a tree is the sequence of the degree of the nodes taken in breadth-first order.

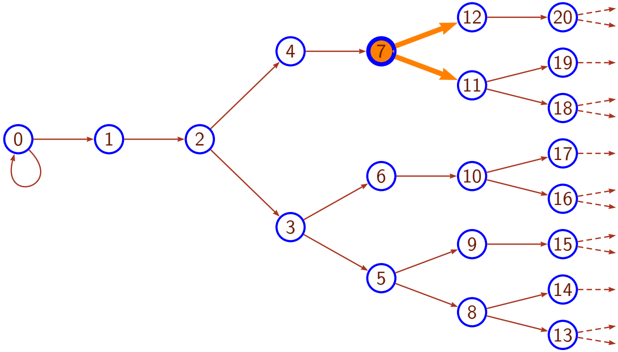


$$\mathbf{s} = 2\ 1\ 2\ 2\ 1\ 2\ 1$$

Signature of a tree

Definition

The **signature** of a tree is the sequence of the degree of the nodes taken in breadth-first order.

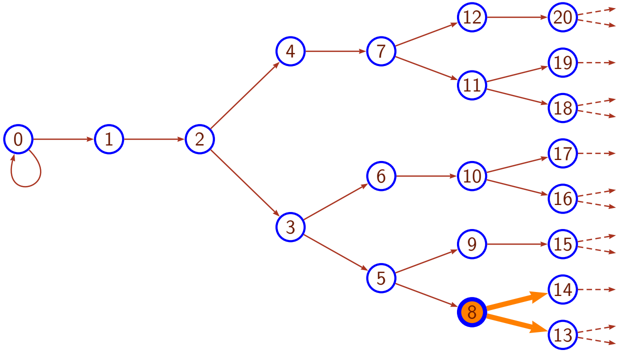


$$\mathbf{s} = 2\ 1\ 2\ 2\ 1\ 2\ 1\ 2$$

Signature of a tree

Definition

The **signature** of a tree is the sequence of the degree of the nodes taken in breadth-first order.

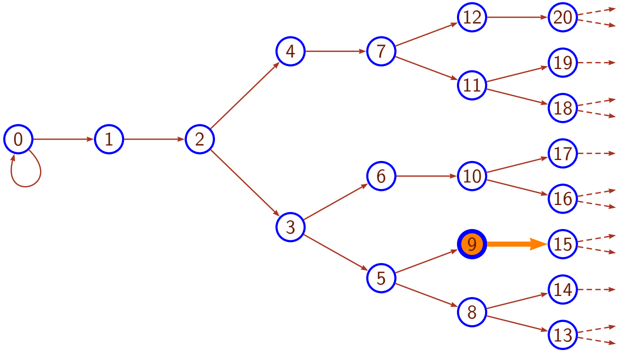


$$\mathbf{s} = 2\ 1\ 2\ 2\ 1\ 2\ 1\ 2\ 2$$

Signature of a tree

Definition

The **signature** of a tree is the sequence of the degree of the nodes taken in breadth-first order.

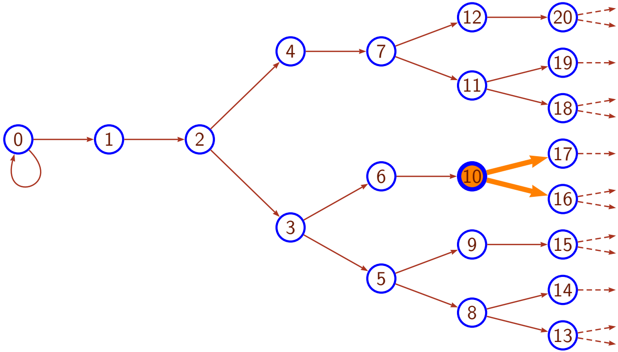


$$\mathbf{s} = 2\ 1\ 2\ 2\ 1\ 2\ 1\ 2\ 2\ 1$$

Signature of a tree

Definition

The **signature** of a tree is the sequence of the degree of the nodes taken in breadth-first order.

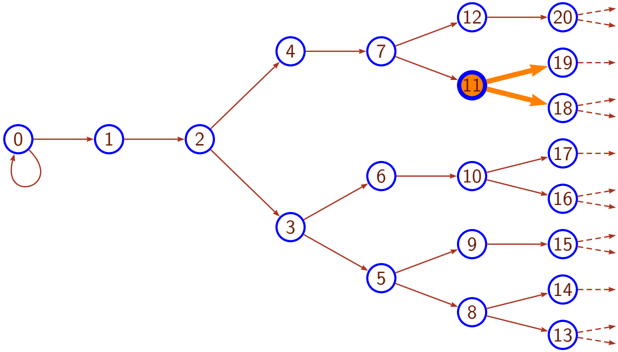


$$\mathbf{s} = 2\ 1\ 2\ 2\ 1\ 2\ 1\ 2\ 2\ 1\ 2$$

Signature of a tree

Definition

The **signature** of a tree is the sequence of the degree of the nodes taken in breadth-first order.

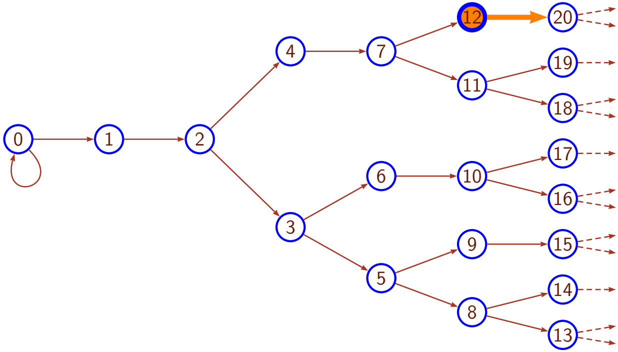


$\mathbf{s} = 2\ 1\ 2\ 2\ 1\ 2\ 1\ 2\ 2\ 1\ 2\ 2$

Signature of a tree

Definition

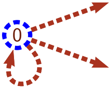
The **signature** of a tree is the sequence of the degree of the nodes taken in breadth-first order.



$$\mathbf{s} = 2\ 1\ 2\ 2\ 1\ 2\ 1\ 2\ 2\ 1\ 2\ 2\ 1\ \cdots$$

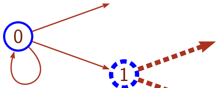
Signature is characteristic of tree

$$\mathbf{s} = (\; 3 \;\; 2 \;\; 1 \;)^{\omega}$$



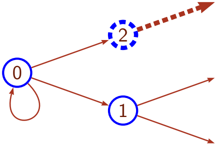
Signature is characteristic of tree

$$\mathbf{s} = (\ 3\ \ 2\ \ 1\ )^{\omega}$$



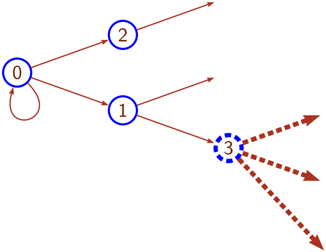
Signature is characteristic of tree

$$\mathbf{s} = (\ 3\ \ 2\ \ \boxed{1}\ )^{\omega}$$



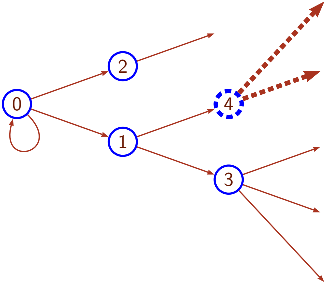
Signature is characteristic of tree

$$\mathbf{s} = (\; 3 \;\; 2 \;\; 1 \;)^{\omega}$$



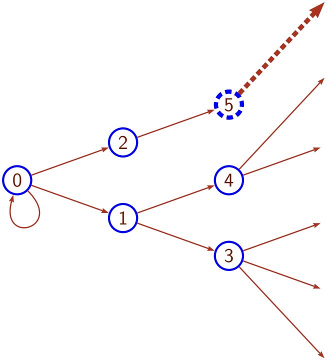
Signature is characteristic of tree

$$\mathbf{s} = (\ 3\ \ 2\ \ 1\ )^{\omega}$$



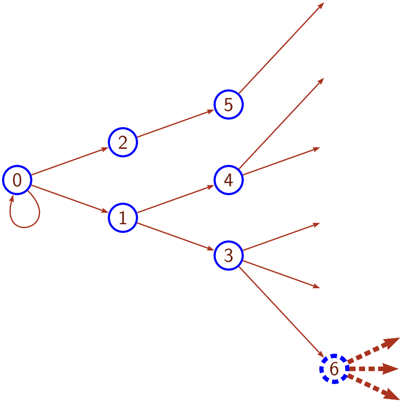
Signature is characteristic of tree

$$\mathbf{s} = (\ 3\ \ 2\ \ \boxed{1}\ )^{\omega}$$



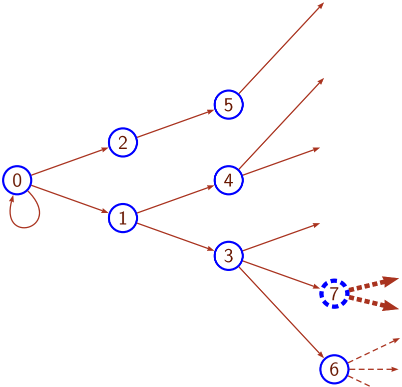
Signature is characteristic of tree

$$\mathbf{s} = (\ 3\ \ 2\ \ 1\ )^{\omega}$$



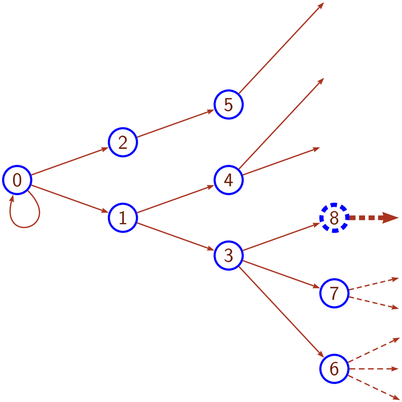
Signature is characteristic of tree

$$\mathbf{s} = (\ 3\ \ 2\ \ 1\ )^{\omega}$$



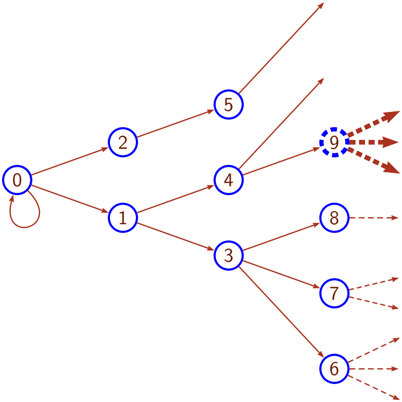
Signature is characteristic of tree

$$\mathbf{s} = (\ 3\ \ 2\ \ \boxed{1}\ )^{\omega}$$



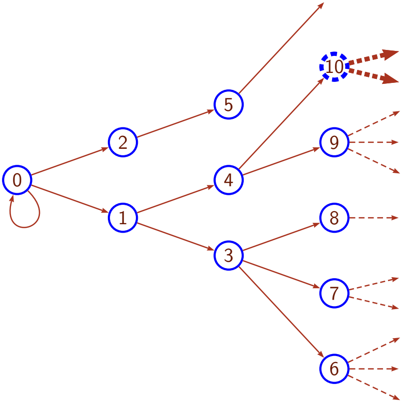
Signature is characteristic of tree

$$\mathbf{s} = (\ 3\ \ 2\ \ 1\ )^{\omega}$$



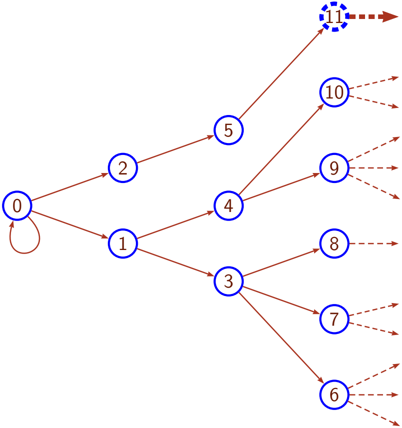
Signature is characteristic of tree

$$\mathbf{s} = (\ 3\ \ 2\ \ 1\ )^{\omega}$$



Signature is characteristic of tree

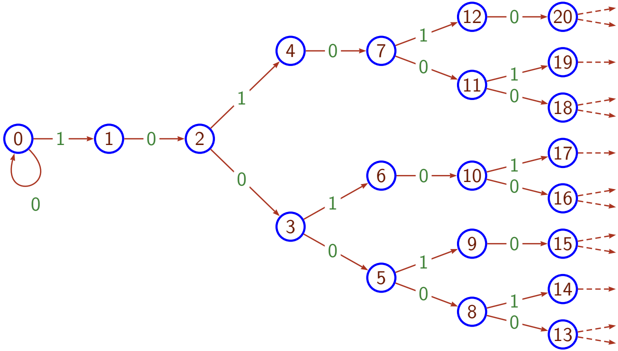
$$\mathbf{s} = (\ 3\ \ 2\ \ 1\ )^{\omega}$$



Prefix-closed languages and labelled trees

Alphabets are ordered hence

prefix-closed languages = labelled trees.



Prefix-closed languages and labelled trees

Alphabets are ordered hence

prefix-closed languages = labelled trees.

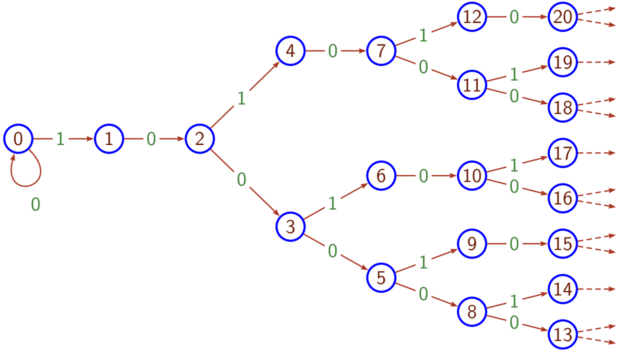


Figure: Integer representations in the Fibonacci numeration system.

Prefix-closed languages and labelled trees

Alphabets are ordered hence

prefix-closed languages = labelled trees.

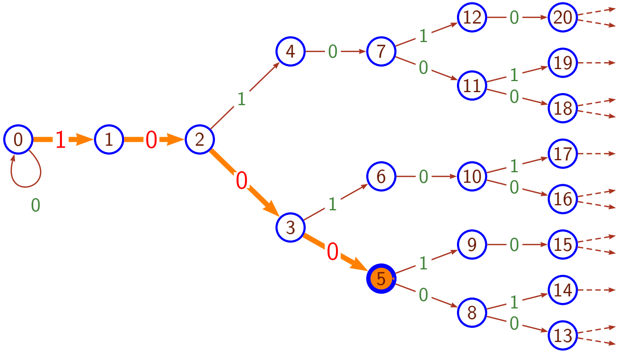


Figure: Integer representations in the Fibonacci numeration system.

Prefix-closed languages and labelled trees

Alphabets are ordered hence

prefix-closed languages = labelled trees.

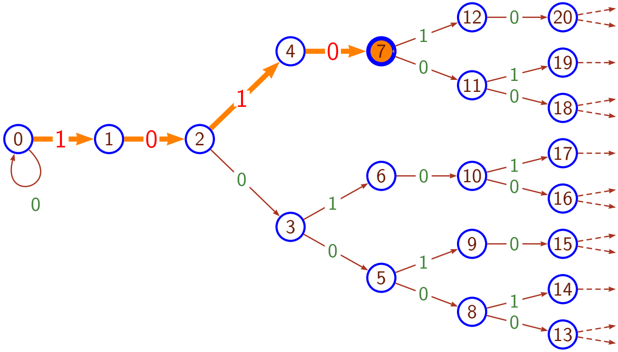
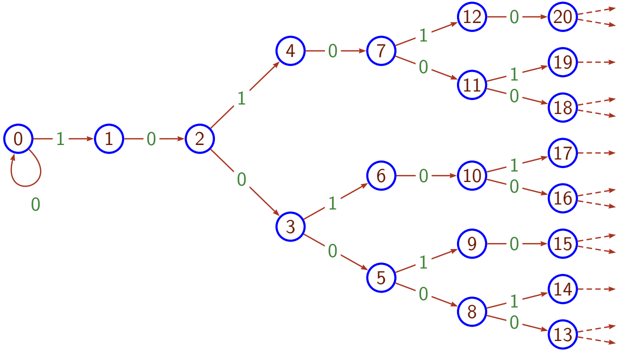


Figure: Integer representations in the Fibonacci numeration system.

Serialisation of a prefix-closed language

Definition

The **labelling** of a language is the **sequence of arc labels** of its transitions taken in breadth-first order.



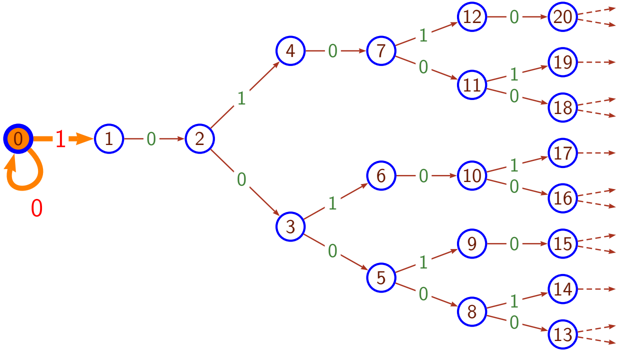
$\mathbf{s} =$

$\lambda =$

Serialisation of a prefix-closed language

Definition

The **labelling** of a language is the **sequence of arc labels** of its transitions taken in breadth-first order.



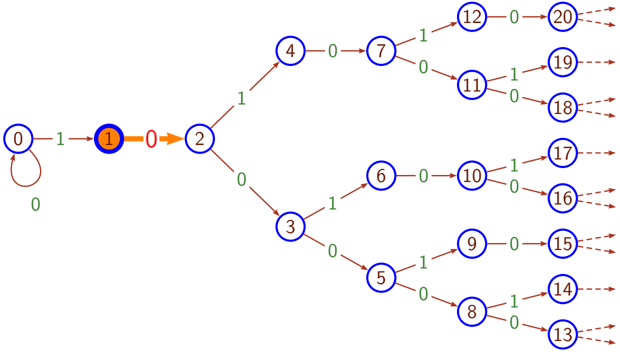
$\mathbf{s} = \boxed{2}$

$\lambda = \boxed{01}$

Serialisation of a prefix-closed language

Definition

The **labelling** of a language is the **sequence of arc labels** of its transitions taken in breadth-first order.



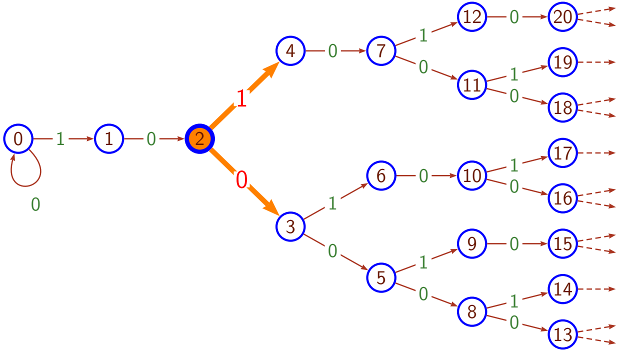
$\mathbf{s} = 2\ 1$

$\lambda = 01\ 0$

Serialisation of a prefix-closed language

Definition

The **labelling** of a language is the **sequence of arc labels** of its transitions taken in breadth-first order.



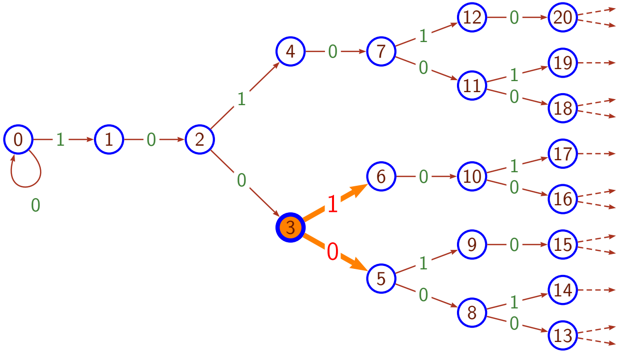
$\mathbf{s} = 2\ \ 1\ \ \boxed{2}$

$\lambda = 01\ 0\ \boxed{01}$

Serialisation of a prefix-closed language

Definition

The **labelling** of a language is the **sequence of arc labels** of its transitions taken in breadth-first order.



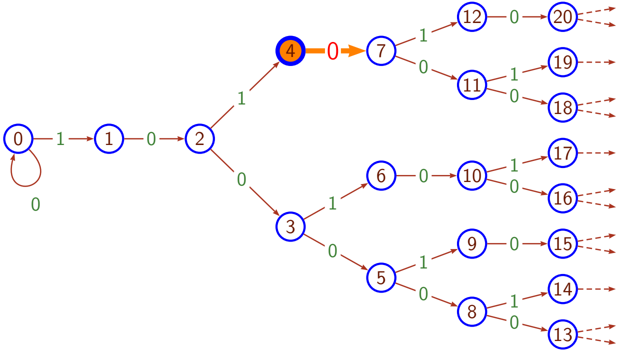
$$\mathbf{s} = 2\ \ 1\ \ 2\ \ \ 2$$

$$\lambda = 01\ 0\ 01\ 01$$

Serialisation of a prefix-closed language

Definition

The **labelling** of a language is the **sequence of arc labels** of its transitions taken in breadth-first order.



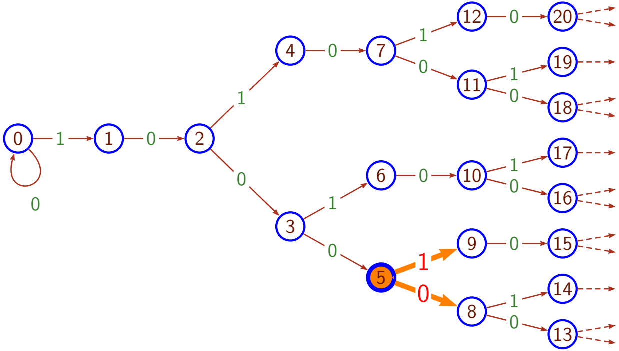
$$\mathbf{s} = 2\ \ 1\ \ 2\ \ \ 2\ \ 1$$

$$\lambda = 01\ 0\ 01\ 01\ 0$$

Serialisation of a prefix-closed language

Definition

The **labelling** of a language is the **sequence of arc labels** of its transitions taken in breadth-first order.



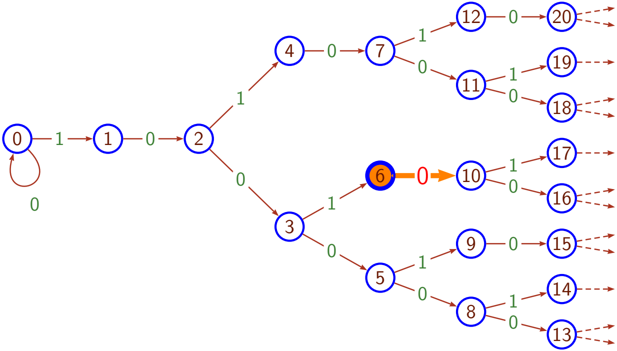
$$\mathbf{s} = 2\ 1\ 2\ \ 2\ 1\ 2$$

$$\lambda = 01\ 0\ 01\ 01\ 0\ 01$$

Serialisation of a prefix-closed language

Definition

The **labelling** of a language is the **sequence of arc labels** of its transitions taken in breadth-first order.



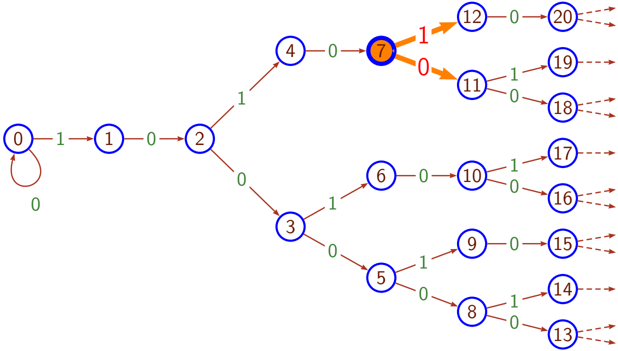
$$\mathbf{s} = 2\;\;1\;\;2\;\;\;2\;\;1\;\;2\;\;1$$

$$\lambda = 01\;0\;01\;01\;0\;01\;0$$

Serialisation of a prefix-closed language

Definition

The **labelling** of a language is the **sequence of arc labels** of its transitions taken in breadth-first order.



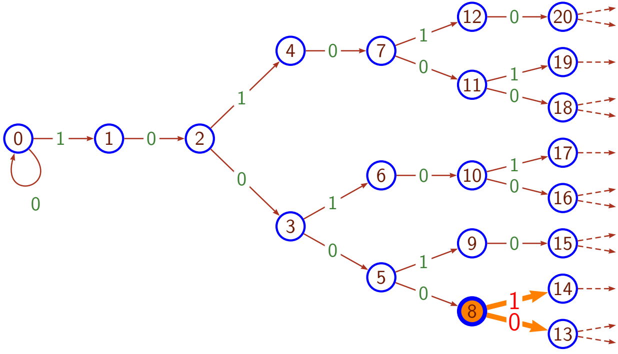
$$\mathbf{s} = 2 \ 1 \ 2 \ \ 2 \ 1 \ 2 \ 1 \ 2$$

$$\lambda = 01 \ 0 \ 01 \ 01 \ 0 \ 01 \ 0 \ 01$$

Serialisation of a prefix-closed language

Definition

The **labelling** of a language is the **sequence of arc labels** of its transitions taken in breadth-first order.



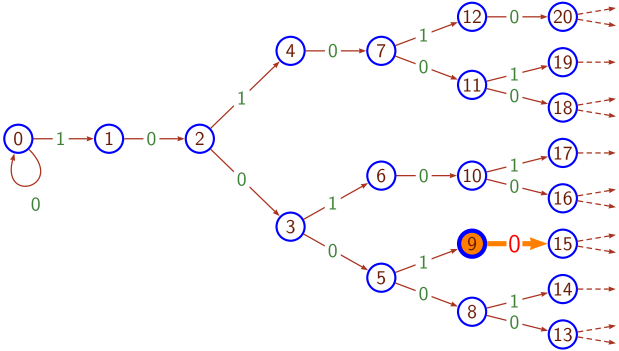
$$\mathbf{s} = 2\ \ 1\ \ 2\ \ \ 2\ \ 1\ \ 2\ \ 1\ \ 2\ \ \ 2$$

$$\lambda = 01\ 0\ 01\ 01\ 0\ 01\ 0\ 01\ 01$$

Serialisation of a prefix-closed language

Definition

The **labelling** of a language is the **sequence of arc labels** of its transitions taken in breadth-first order.



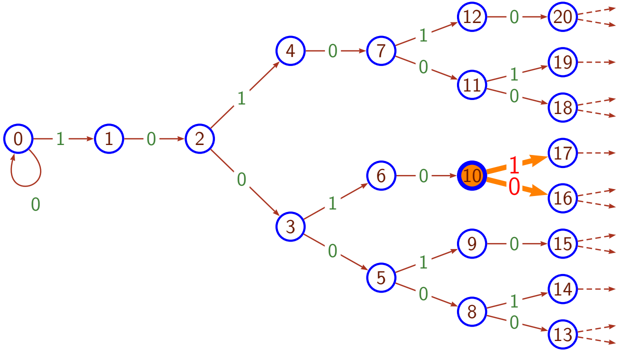
$$\mathbf{s} = 2\ 1\ 2\ \ 2\ 1\ 2\ 1\ 2\ \ 2\ 1$$

$$\lambda = 01\ 0\ 01\ 01\ 0\ 01\ 0\ 01\ 01\ 0$$

Serialisation of a prefix-closed language

Definition

The **labelling** of a language is the **sequence of arc labels** of its transitions taken in breadth-first order.



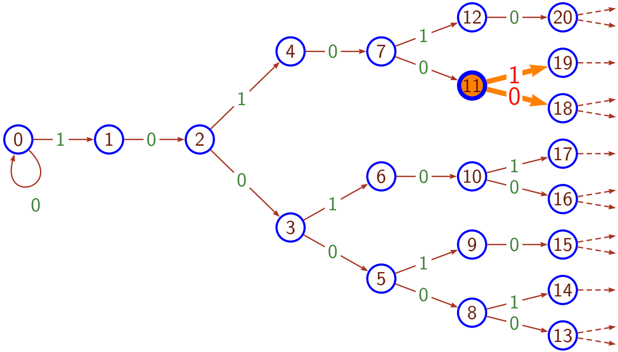
$$\mathbf{s} = 2\ 1\ 2\ \ 2\ 1\ 2\ 1\ 2\ \ 2\ 1\ 2$$

$$\lambda = 01\ 0\ 01\ 01\ 0\ 01\ 0\ 01\ 01\ 0\ 01$$

Serialisation of a prefix-closed language

Definition

The **labelling** of a language is the **sequence of arc labels** of its transitions taken in breadth-first order.



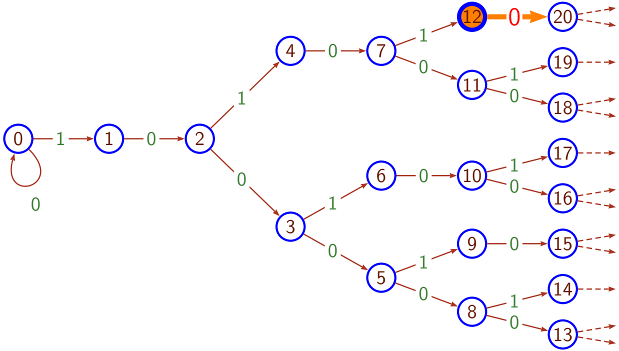
$$\mathbf{s} = 2\ 1\ 2\ \ 2\ 1\ 2\ 1\ 2\ \ 2\ 1\ 2\ \ 2$$

$$\lambda = 01\ 0\ 01\ 01\ 0\ 01\ 0\ 01\ 01\ 0\ 01\ 01$$

Serialisation of a prefix-closed language

Definition

The **labelling** of a language is the **sequence of arc labels** of its transitions taken in breadth-first order.

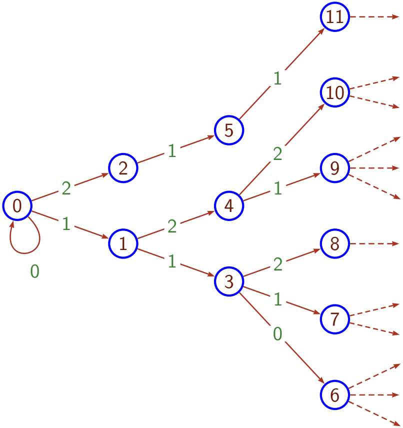


$$\mathbf{s} = 2\ 1\ 2\ \ 2\ 1\ 2\ 1\ 2\ \ 2\ 1\ 2\ \ 2\ 1\ \cdots$$

$$\lambda = 01\ 0\ 01\ 01\ 0\ 01\ 0\ 01\ 01\ 0\ 01\ 01\ 0\ \cdots$$

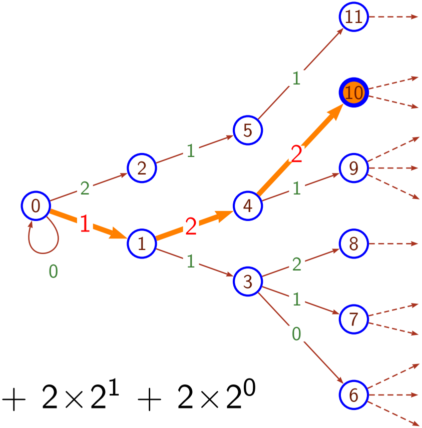
The pair signature/labelling is characteristic

$$\mathbf{s} = (3\ 2\ 1)^{\omega}$$
$$\lambda = (012\ 12\ 1)^{\omega}$$



The pair signature/labelling is characteristic

$$\mathbf{s} = (3\ 2\ 1)^\omega$$
$$\lambda = (012\ 12\ 1)^\omega$$



$$10 = 1 \times 2^2 + 2 \times 2^1 + 2 \times 2^0$$

Figure: Non-canonical integer representations in base 2.

Abstract Numeration System (Lecomte-Rigo)

Observation

In basically every NS, the representations of integers follows the *radix order*: $\forall n, p \quad \langle n \rangle \leq_{rad} \langle n+p \rangle$

Abstract Numeration System (Lecomte-Rigo)

Observation

In basically every NS, the representations of integers follows the *radix order*: $\forall n, p \quad \langle n \rangle \leq_{rad} \langle n+p \rangle$

$$\begin{array}{lll} u <_{\text{rad}} v & \text{if} & |u| < |v| \\ & \text{or} & |u| = |v| \ \ \& \ \ u <_{\text{lex}} v \end{array}$$

Example: $2 <_{\text{rad}} 12 \qquad 12 <_{\text{rad}} 21$.

Abstract Numeration System (Lecomte-Rigo)

Observation

In basically every NS, the representations of integers follows the *radix order*: $\forall n, p \quad \langle n \rangle \leq_{rad} \langle n+p \rangle$

$$u <_{\text{rad}} v \quad \text{if} \quad |u| < |v| \quad \text{or} \quad |u| = |v| \;\&\; u <_{\text{lex}} v$$

Example: $2 <_{\text{rad}} 12 \qquad 12 <_{\text{rad}} 21$.

Definition (ANS L)

L: language over an ordered alphabet A.
$\langle n \rangle_L$ is the $(n+1)$-th word of L in the radix order.

In our scheme, $\langle n \rangle_L$ is the word labelling the path $0 \longrightarrow n$.

Examples of ANS's

$L = \{ \ u \in \{0, 1, \ldots, p-1\}^* \mid u$ does not start with a 0 $\}$
→ NS in base p

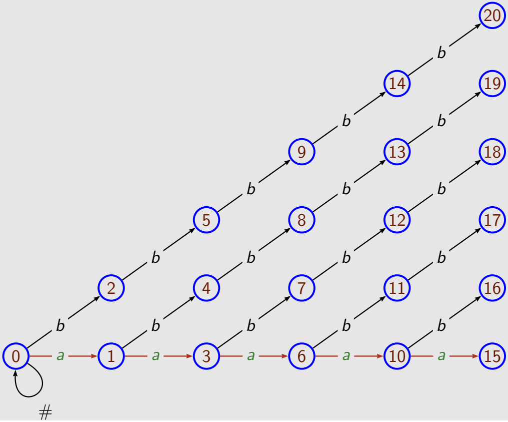
$L = \{ \ u \in \{0, 1\}^* \mid u$ does not contain the factor 11 $\}$
→ Fibonacci NS

Examples of ANS's

$L = \{\ u \in \{0, 1, \ldots, p-1\}^* \mid u$ does not start with a 0 $\}$
$\to$ NS in base p

$L = \{\ u \in \{0, 1\}^* \mid u$ does not contain the factor 11 $\}$
$\to$ Fibonacci NS

$L = a^* b^*$

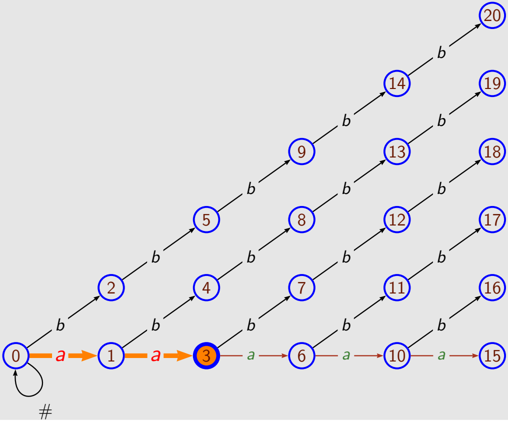


Examples of ANS's

$L = \{\ u \in \{0, 1, \ldots, p-1\}^* \mid u \text{ does not start with a } 0\ \}$
→ NS in base p

$L = \{\ u \in \{0, 1\}^* \mid u \text{ does not contain the factor } 11\ \}$
→ Fibonacci NS

$L = a^* b^*$
$\langle 3 \rangle = aa$

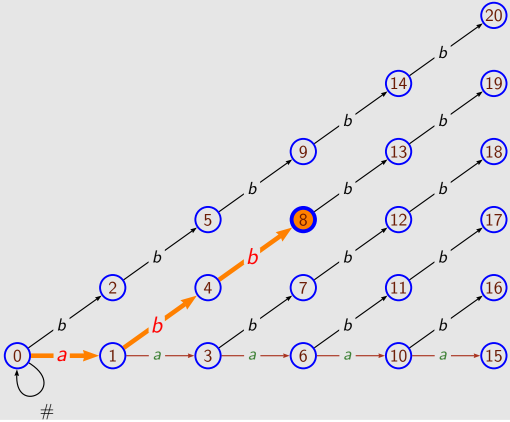


Examples of ANS's

$L = \{\ u \in \{0, 1, \ldots, p-1\}^* \mid u$ does not start with a 0 $\}$
→ NS in base p

$L = \{\ u \in \{0, 1\}^* \mid u$ does not contain the factor 11 $\}$
→ Fibonacci NS

$L = a^* b^*$
$\langle 3 \rangle = aa$
$\langle 8 \rangle = abb$



Outline

NS = Numeration system

Prefix-closed Abstract Rational NS (Lecomte–Rigo 2001)

Built from an arbitrary prefix-closed regular language.

Dumont-Thomas NS (Dumont-Thomas, 1989)

Built from an arbitrary morphism.

NS = Numeration system

Prefix-closed Abstract Rational NS (Lecomte–Rigo 2001)

Built from an arbitrary prefix-closed regular language.

Dumont-Thomas NS (Dumont-Thomas, 1989)

Built from an arbitrary morphism.

Theorem

L: a prefix-closed language.
Signature(L) is a morphic word $\Leftrightarrow$ L is a regular language.

NS = Numeration system

Prefix-closed Abstract Rational NS (Lecomte–Rigo 2001)

Built from an arbitrary prefix-closed regular language.

Dumont-Thomas NS (Dumont-Thomas, 1989)

Built from an arbitrary morphism.

Theorem

L: a prefix-closed language.
Signature(L) is a morphic word $\Leftrightarrow$ L is a regular language.

Theorem

Every DTNS is a prefix-closed ARNS.

Every prefix-closed ARNS is easily[†] convertible to a DTNS.

[†] Through a finite, letter-to-letter and pure sequential transducer.

Theorem

L: a prefix-closed language. Signature(L) is a morphic signature $\Leftrightarrow$ L is a regular language.

Word morphisms

σ: a morphism $A^* \to A^*$.

Running examples

Fibonacci morphism: $\{a, b\} \to \{a, b\}^*$

$a \mapsto ab$

$b \mapsto a$

Word morphisms

σ: a morphism $A^* \to A^*$.

Running examples

Fibonacci morphism: $\{a, b\} \to \{a, b\}^*$

$a \mapsto ab$

$b \mapsto a$

A periodic morphism: $\{a, b, c\} \to \{a, b, c\}^*$

$a \mapsto abc$

$b \mapsto ab$

$c \mapsto c$

Word morphisms

σ: a morphism $A^* \to A^*$.

σ is prolongable on a if $\sigma(a)$ starts with the letter a.

Running examples

Fibonacci morphism: $\{a, b\} \to \{a, b\}^*$

$a \mapsto ab$

$b \mapsto a$

A periodic morphism: $\{a, b, c\} \to \{a, b, c\}^*$

$a \mapsto abc$

$b \mapsto ab$

$c \mapsto c$

Word morphisms

σ: a morphism $A^* \to A^*$.

σ is prolongable on a if $\sigma(a)$ starts with the letter a.
In this case, $\sigma^\omega(a)$ exists and is called a pure morphic word.

Running examples

Fibonacci morphism: $\{a, b\} \to \{a, b\}^*$
$a \mapsto ab$
$b \mapsto a$

A periodic morphism: $\{a, b, c\} \to \{a, b, c\}^*$
$a \mapsto abc$
$b \mapsto ab$
$c \mapsto c$

Word morphisms

σ: a morphism $A^* \to A^*$.

σ is prolongable on a if $\sigma(a)$ starts with the letter a.
In this case, $\sigma^\omega(a)$ exists and is called a pure morphic word.

f : a letter-to-letter morphism $A^* \to B^*$.
→ $f(\sigma^\omega(a))$ is called a morphic word.

Running examples

Fibonacci morphism: $\{a, b\} \to \{a, b\}^*$
$a \mapsto ab$
$b \mapsto a$

A periodic morphism: $\{a, b, c\} \to \{a, b, c\}^*$
$a \mapsto abc$
$b \mapsto ab$
$c \mapsto c$

Morphic signature

let $f_\sigma : A^* \to D^*$ be the (letter-to-letter) morphism defined by

- $D \subset \mathbb{N}$
- $\forall b,\ f_\sigma(b) = |\sigma(b)|$

We call $f_\sigma(\sigma^\omega(a))$ a **morphic signature**.

Example: Fibonacci morphism

$\sigma(a) = ab$

$\sigma(b) = a$

Morphic signature

let $f_\sigma : A^* \to D^*$ be the (letter-to-letter) morphism defined by

- $D \subset \mathbb{N}$
- $\forall b,\ f_\sigma(b) = |\sigma(b)|$

We call $f_\sigma(\sigma^\omega(a))$ a **morphic signature**.

Example: Fibonacci morphism

$\sigma(a) = ab \quad \Longrightarrow \quad f_\sigma(a) = 2$

$\sigma(b) = a \quad \Longrightarrow \quad f_\sigma(b) = 1$

Morphic signature

let $f_\sigma : A^* \to D^*$ be the (letter-to-letter) morphism defined by

- $D \subset \mathbb{N}$
- $\forall b,\ f_\sigma(b) = |\sigma(b)|$

We call $f_\sigma(\sigma^\omega(a))$ a **morphic signature**.

Example: Fibonacci morphism

$$\sigma(a) = ab \quad \Longrightarrow \quad f_\sigma(a) = 2$$

$$\sigma(b) = a \quad \Longrightarrow \quad f_\sigma(b) = 1$$

$$f_\sigma(\sigma^\omega(a)) \quad = \quad 2122121221221212212122\cdots$$

Morphic signature

let $f_\sigma : A^* \to D^*$ be the (letter-to-letter) morphism defined by

- $D \subset \mathbb{N}$
- $\forall b,\ f_\sigma(b) = |\sigma(b)|$

We call $f_\sigma(\sigma^\omega(a))$ a **morphic signature**.

Example: a periodic morphism

$\sigma(a) = abc$

$\sigma(b) = ab$

$\sigma(c) = c$

Morphic signature

let $f_\sigma : A^* \to D^*$ be the (letter-to-letter) morphism defined by

- $D \subset \mathbb{N}$
- $\forall b,\ f_\sigma(b) = |\sigma(b)|$

We call $f_\sigma(\sigma^\omega(a))$ a **morphic signature**.

Example: a periodic morphism

$\sigma(a) = abc \quad \Longrightarrow f_\sigma(a) = 3$

$\sigma(b) = ab \quad \Longrightarrow f_\sigma(b) = 2$

$\sigma(c) = c \quad \Longrightarrow f_\sigma(c) = 1$

Morphic signature

let $f_\sigma : A^* \to D^*$ be the (letter-to-letter) morphism defined by

- $D \subset \mathbb{N}$
- $\forall b,\ f_\sigma(b) = |\sigma(b)|$

We call $f_\sigma(\sigma^\omega(a))$ a **morphic signature**.

Example: a periodic morphism

$$\sigma(a) = abc \quad \Longrightarrow f_\sigma(a) = 3$$

$$\sigma(b) = ab \quad \Longrightarrow f_\sigma(b) = 2$$

$$\sigma(c) = c \quad \Longrightarrow f_\sigma(c) = 1$$

$$\sigma(abc) \quad = \quad abc\,abc \qquad \text{hence} \quad f_\sigma(\sigma^\omega(a)) \quad = \quad (321)^\omega$$

Morphic labelling

If g is a morphism such that

- $\forall b,\ |g(b)| = |\sigma(b)|$
- if $g(b) = c_0 c_1 \cdots c_k$ then $c_0 < c_1 < \cdots < c_k$

We call $g(\sigma^\omega(a))$ a **morphic labelling**.

Morphic labelling

If g is a morphism such that

- $\forall b,\ |g(b)| = |\sigma(b)|$
- if $g(b) = c_0 c_1 \cdots c_k$ then $c_0 < c_1 < \cdots < c_k$

We call $g(\sigma^\omega(a))$ a **morphic labelling**.

Example: Fibonacci morphism

$$\sigma(a) = \text{ab} \implies f_\sigma(a) = 2$$
$$\sigma(b) = \text{a} \implies f_\sigma(b) = 1$$
$$f_\sigma(\sigma^\omega(a)) = 2122121221221212212122\cdots$$

If we choose g:

$g(a) = 01$

$g(b) = 0$

Morphic labelling

If g is a morphism such that

- $\forall b,\ |g(b)| = |\sigma(b)|$
- if $g(b) = c_0 c_1 \cdots c_k$ then $c_0 < c_1 < \cdots < c_k$

We call $g(\sigma^\omega(a))$ a **morphic labelling**.

Example: Fibonacci morphism

$$\sigma(a) = \mathrm{ab} \quad \Longrightarrow \quad f_\sigma(a) = 2$$
$$\sigma(b) = \mathrm{a} \quad \Longrightarrow \quad f_\sigma(b) = 1$$
$$f_\sigma(\sigma^\omega(a)) \quad = \quad 2122121221221212212122\cdots$$

If we choose g:

$g(a) = 01$

$g(b) = 0$

Morphic labelling

If g is a morphism such that

- $\forall b, \ |g(b)| = |\sigma(b)|$
- if $g(b) = c_0 c_1 \cdots c_k$ then $c_0 < c_1 < \cdots < c_k$

We call $g(\sigma^\omega(a))$ a **morphic labelling**.

Example: Fibonacci morphism

$$\begin{aligned} \sigma(a) &= \text{ab} & \implies & \ f_\sigma(a) = 2 \\ \sigma(b) &= \text{a} & \implies & \ f_\sigma(b) = 1 \end{aligned}$$

$$f_\sigma(\sigma^\omega(a)) \quad = \quad 2122121221221212212122\cdots$$

If we choose g:

$g(a) = 01$

$g(b) = 0$

Morphic labelling

If g is a morphism such that

- $\forall b,\ |g(b)| = |\sigma(b)|$
- if $g(b) = c_0 c_1 \cdots c_k$ then $c_0 < c_1 < \cdots < c_k$

We call $g(\sigma^\omega(a))$ a **morphic labelling**.

Example: Fibonacci morphism

$$\begin{aligned} \sigma(a) &= \text{ab} & \Longrightarrow\ & f_\sigma(a) = 2 \\ \sigma(b) &= \text{a} & \Longrightarrow\ & f_\sigma(b) = 1 \end{aligned}$$

$$f_\sigma(\sigma^\omega(a)) \quad = \quad 2122121221221212212122\cdots$$

If we choose g:

$$g(a) = 01$$

$$g(b) = 0$$

$$g(\sigma^\omega(a)) \quad = \quad 01\,0\,01\,01\,0\,01\,0\,01\,01\,0\,01\,01\,0\ \cdots$$

Morphic labelling

If g is a morphism such that

- $\forall b,\ |g(b)| = |\sigma(b)|$
- if $g(b) = c_0 c_1 \cdots c_k$ then $c_0 < c_1 < \cdots < c_k$

We call $g(\sigma^\omega(a))$ a **morphic labelling**.

Example: a periodic morphism

$$\sigma(a) = abc \quad \Longrightarrow f_\sigma(a) = 3$$
$$\sigma(b) = ab \quad \Longrightarrow f_\sigma(b) = 2$$
$$\sigma(c) = c \quad \Longrightarrow f_\sigma(c) = 1$$
$$\sigma(abc) \quad = \quad abc\,abc \qquad \text{hence } f_\sigma(\sigma^\omega(a)) \quad = \quad (321)^\omega$$

If we choose g:

$g(a) = 012$

$g(b) = 12$

$g(c) = 1$

$$g(\sigma^\omega(a)) \quad = \quad (012\,12\,1)^\omega$$

Forward direction of the theorem

Theorem

L: a prefix-closed language.

$$\text{Signature}(L) \text{ is morphic} \quad \Leftrightarrow \quad L \text{ is a regular language.}$$

Forward direction of the theorem

Theorem

L: a prefix-closed language.

Signature(L) is morphic $\Leftrightarrow$ L is a regular language.

(σ, g): a substitutive signature.
(σ, g) defines a finite automaton $\mathcal{A}_{(\sigma,g)}$.
It is analogous to

- the *prefix graph/automaton* in Dumont–Thomas '89,'91,'93
- or the correspondence used in Maes–Rigo '02.

Forward direction of the theorem

Theorem

L: a prefix-closed language.

Signature(L) is morphic $\Leftrightarrow$ L is a regular language.

(σ, g): a substitutive signature.
(σ, g) defines a finite automaton $\mathcal{A}_{(\sigma,g)}$.
It is analogous to

- the *prefix graph/automaton* in Dumont–Thomas '89,'91,'93
- or the correspondence used in Maes–Rigo '02.

Proposition

The language accepted by $\mathcal{A}_{(\sigma,g)}$ has signature (σ, g).

Automaton associated with a morphic signature

$\sigma : \mathsf{A}^* \to \mathsf{A}^*$ prolongable on a and $g : \mathsf{A}^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle\, \mathsf{A} \,, B, \, \delta \,, \, \{\mathsf{a}\} \,, \, \mathsf{A} \,\rangle$$

$\sigma(\,\mathsf{a}\,) = \mathsf{a}\,\mathsf{b}$
$\sigma(\,\mathsf{b}\,) = \mathsf{a}$

$g(\,\mathsf{a}\,) = 0\,1$
$g(\,\mathsf{b}\,) = 0$

Automaton associated with a morphic signature

$\sigma : \mathsf{A}^* \to \mathsf{A}^*$ prolongable on a and $g : \mathsf{A}^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle \mathsf{A}, B, \delta, \{\mathsf{a}\}, \mathsf{A} \rangle$$

$\sigma(\mathsf{a}) = \mathsf{a}\,\mathsf{b}$
$\sigma(\mathsf{b}) = \mathsf{a}$

$g(\mathsf{a}) = 0\,1$
$g(\mathsf{b}) = 0$

Automaton associated with a morphic signature

$\sigma : \mathsf{A}^* \to \mathsf{A}^*$ prolongable on a and $g : \mathsf{A}^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle\, \mathsf{A} \,, B,\, \delta \,,\, \{\mathsf{a}\},\, \mathsf{A} \,\rangle$$

$\sigma(\,\mathsf{a}\,) = \mathsf{a}\,\mathsf{b}$
$\sigma(\,\mathsf{b}\,) = \mathsf{a}$

$g(\,\mathsf{a}\,) = 0\,1$
$g(\,\mathsf{b}\,) = 0$

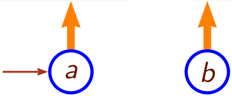
Automaton associated with a morphic signature

$\sigma : \mathrm{A}^* \to \mathrm{A}^*$ prolongable on a and $g : \mathrm{A}^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle\, \mathrm{A}\,, B, \,\delta\,,\, \{\mathrm{a}\}\,,\, \mathrm{A}\,\rangle$$

$\sigma(\mathrm{a}) = \mathrm{a\,b}$
$\sigma(\mathrm{b}) = \mathrm{a}$

$g(\mathrm{a}) = 0\,1$
$g(\mathrm{b}) = 0$



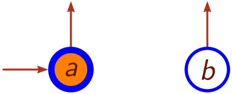
Automaton associated with a morphic signature

$\sigma : \mathsf{A}^* \to \mathsf{A}^*$ prolongable on a and $g : \mathsf{A}^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle\, \mathsf{A}\,, B, \delta, \{\mathsf{a}\}\,, \mathsf{A} \,\rangle$$

$\sigma(\,\mathsf{a}\,) = \mathsf{a}\,\mathsf{b}$
$\sigma(\,\mathsf{b}\,) = \mathsf{a}$

$g(\,\mathsf{a}\,) = 0\,1$
$g(\,\mathsf{b}\,) = 0$



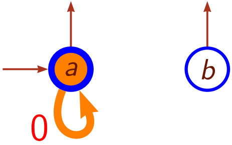
Automaton associated with a morphic signature

$\sigma : \mathsf{A}^* \to \mathsf{A}^*$ prolongable on a and $g : \mathsf{A}^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle\, \mathsf{A}\,, B, \delta, \{\mathsf{a}\}\,, \mathsf{A}\,\rangle$$

$\sigma(\,\mathsf{a}\,) = \mathsf{a}\,\mathsf{b}$
$\sigma(\,\mathsf{b}\,) = \mathsf{a}$

$g(\,\mathsf{a}\,) = 0\,1$
$g(\,\mathsf{b}\,) = 0$



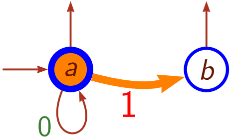
Automaton associated with a morphic signature

$\sigma : \mathsf{A}^* \to \mathsf{A}^*$ prolongable on a and $g : \mathsf{A}^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle\, \mathsf{A}\,, B, \delta, \{\mathsf{a}\}\,, \mathsf{A}\,\rangle$$

$\sigma(\,\mathsf{a}\,) = \mathsf{a}\,\mathsf{b}$
$\sigma(\,\mathsf{b}\,) = \mathsf{a}$

$g(\,\mathsf{a}\,) = 0\,1$
$g(\,\mathsf{b}\,) = 0$



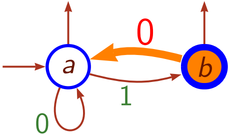
Automaton associated with a morphic signature

$\sigma : \mathsf{A}^* \to \mathsf{A}^*$ prolongable on a and $g : \mathsf{A}^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle\, \mathsf{A} \,, B, \delta, \{\mathsf{a}\} \,, \mathsf{A} \,\rangle$$

$\sigma(\mathsf{a}) = \mathsf{a}\,\mathsf{b}$
$\sigma(\mathsf{b}) = \mathsf{a}$

$g(\mathsf{a}) = 0\,1$
$g(\mathsf{b}) = 0$



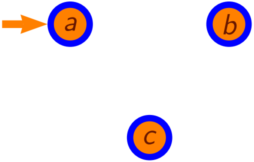
Automaton associated with a morphic signature

$\sigma : \mathsf{A}^* \to \mathsf{A}^*$ prolongable on a and $g : \mathsf{A}^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle \mathsf{A}, B, \delta, \{\mathsf{a}\}, \mathsf{A} \rangle$$

$\sigma(\mathsf{a}) = \mathsf{a\,b\,c}$
$\sigma(\mathsf{b}) = \mathsf{a\,b}$
$\sigma(\mathsf{c}) = \mathsf{c}$

$g(\mathsf{a}) = 0\,1\,2$
$g(\mathsf{b}) = 1\,2$
$g(\mathsf{c}) = 1$



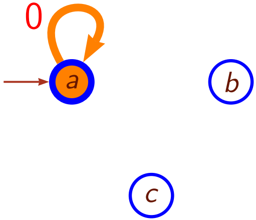
Automaton associated with a morphic signature

$\sigma : A^* \to A^*$ prolongable on a and $g : A^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle\, A\,, B, \delta, \{a\}\,, A\,\rangle$$

$\sigma(\, a\,) = a\,b\,c$
$\sigma(\, b\,) = a\,b$
$\sigma(\, c\,) = c$

$g(\, a\,) = 0\,1\,2$
$g(\, b\,) = 1\,2$
$g(\, c\,) = 1$



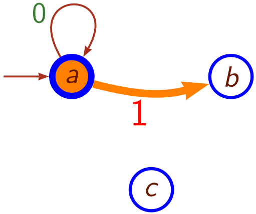
Automaton associated with a morphic signature

$\sigma\ :\ \mathsf{A}^* \to \mathsf{A}^*$ prolongable on a and $g\ :\ \mathsf{A}^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle\, \mathsf{A}\,, B, \delta, \{\mathsf{a}\}\,, \mathsf{A}\,\rangle$$

$\sigma(\,\mathsf{a}\,) = \mathsf{a\,b\,c}$
$\sigma(\,\mathsf{b}\,) = \mathsf{a\,b}$
$\sigma(\,\mathsf{c}\,) = \mathsf{c}$

$g(\,\mathsf{a}\,) = 0\,1\,2$
$g(\,\mathsf{b}\,) = 1\,2$
$g(\,\mathsf{c}\,) = 1$



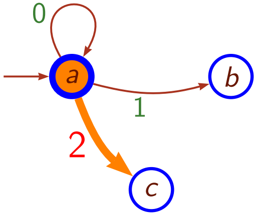
Automaton associated with a morphic signature

$\sigma : \mathsf{A}^* \to \mathsf{A}^*$ prolongable on a and $g : \mathsf{A}^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle\, \mathsf{A}\,, B, \delta, \{\mathsf{a}\}\,, \mathsf{A}\,\rangle$$

$\sigma(\,\mathsf{a}\,) = \mathsf{a\,b\,c}$
$\sigma(\,\mathsf{b}\,) = \mathsf{a\,b}$
$\sigma(\,\mathsf{c}\,) = \mathsf{c}$

$g(\,\mathsf{a}\,) = 0\,1\,2$
$g(\,\mathsf{b}\,) = 1\,2$
$g(\,\mathsf{c}\,) = 1$



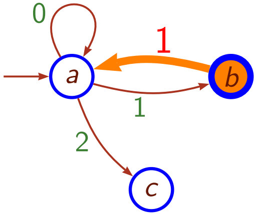
Automaton associated with a morphic signature

$\sigma : \mathsf{A}^* \to \mathsf{A}^*$ prolongable on a and $g : \mathsf{A}^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle\, \mathsf{A}\,, B, \delta, \{\mathsf{a}\}\,, \mathsf{A}\,\rangle$$

$\sigma(\mathsf{a}) = \mathsf{a\,b\,c}$
$\sigma(\mathsf{b}) = \mathsf{a\,b}$
$\sigma(\mathsf{c}) = \mathsf{c}$

$g(\mathsf{a}) = 0\,1\,2$
$g(\mathsf{b}) = 1\,2$
$g(\mathsf{c}) = 1$



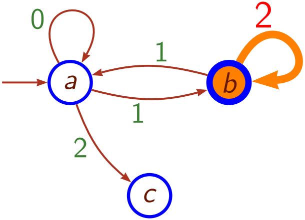
Automaton associated with a morphic signature

$\sigma : \mathsf{A}^* \to \mathsf{A}^*$ prolongable on a and $g : \mathsf{A}^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle\, \mathsf{A}\,, B, \delta, \{\mathsf{a}\}\,, \mathsf{A}\,\rangle$$

$\sigma(\,\mathsf{a}\,) = \mathsf{a\,b\,c}$
$\sigma(\,\mathsf{b}\,) = \mathsf{a\,b}$
$\sigma(\,\mathsf{c}\,) = \mathsf{c}$

$g(\,\mathsf{a}\,) = 0\,1\,2$
$g(\,\mathsf{b}\,) = 1\,2$
$g(\,\mathsf{c}\,) = 1$



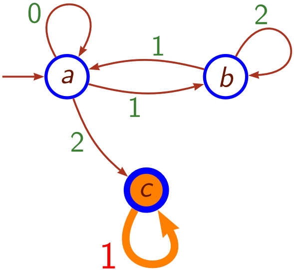
Automaton associated with a morphic signature

$\sigma : \mathsf{A}^* \to \mathsf{A}^*$ prolongable on a and $g : \mathsf{A}^* \to B^*$

$$\mathcal{A}_{(\sigma,g)} = \langle\, \mathsf{A} \,, B, \delta, \{\mathsf{a}\} \,, \mathsf{A} \,\rangle$$

$\sigma(\mathsf{a}) = \mathsf{a\,b\,c}$
$\sigma(\mathsf{b}) = \mathsf{a\,b}$
$\sigma(\mathsf{c}) = \mathsf{c}$

$g(\mathsf{a}) = 0\,1\,2$
$g(\mathsf{b}) = 1\,2$
$g(\mathsf{c}) = 1$



Forward direction of the theorem

Theorem

L: a prefix-closed language.
Signature(L) is morphic $\Leftrightarrow$ L is a regular language.

(σ, g): a substitutive signature.
(σ, g) defines a finite automaton $\mathcal{A}_{(\sigma,g)}$.
It is analogous to

- the *prefix graph/automaton* in Dumont–Thomas '89,'91,'93
- or the correspondence used in Maes–Rigo '02.

Proposition

The language accepted by $\mathcal{A}_{(\sigma,g)}$ has signature (σ, g).

Forward direction of the theorem

Theorem

L: a prefix-closed language.

Signature(L) is morphic $\quad\Leftrightarrow\quad$ L is a regular language.

(σ, g): a substitutive signature.
(σ, g) defines a finite automaton $\mathcal{A}_{(\sigma,g)}$.
It is analogous to

- the *prefix graph/automaton* in Dumont–Thomas '89,'91,'93
- or the correspondence used in Maes–Rigo '02.

Proposition

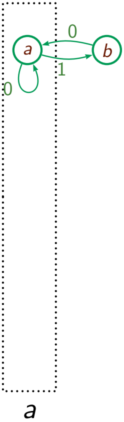
The language accepted by $\mathcal{A}_{(\sigma,g)}$ has signature (σ, g).

Idea of proof: Unfold the automaton $\mathcal{A}_{(\sigma,g)}$.

Idea of proof: unfolding the automaton $\mathcal{A}_{(\sigma,g)}$

$$\sigma(a) = ab$$
$$\sigma(b) = a$$

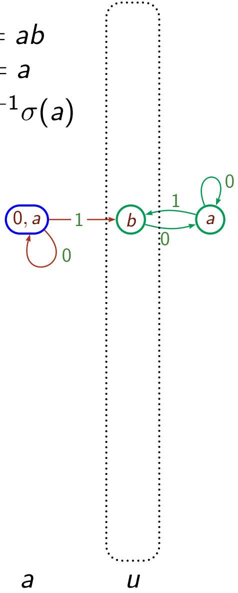
$$u = a^{-1}\sigma(a) = b$$



Idea of proof: unfolding the automaton $\mathcal{A}_{(\sigma,g)}$

$\sigma(a) = ab$
$\sigma(b) = a$

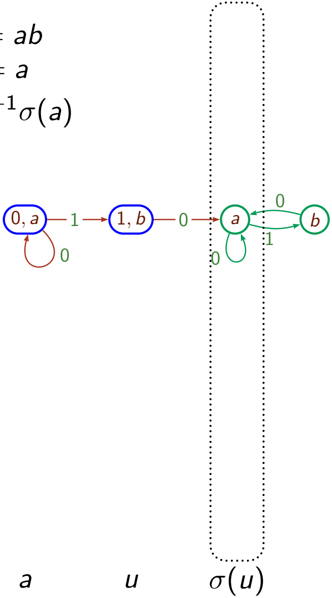
$u = a^{-1}\sigma(a)$
$\quad = b$



Idea of proof: unfolding the automaton $\mathcal{A}_{(\sigma,g)}$

$\sigma(a) = ab$
$\sigma(b) = a$

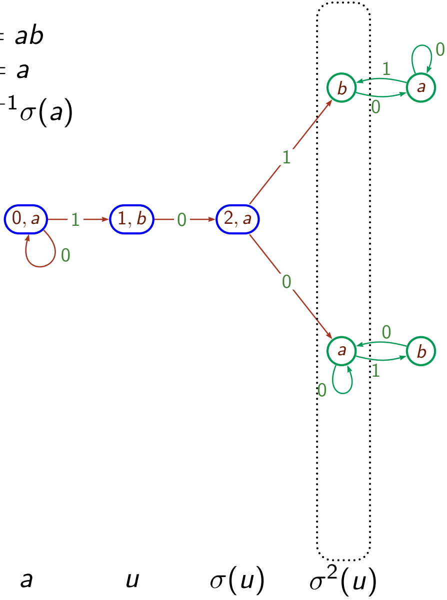
$u = a^{-1}\sigma(a)$
$= b$



Idea of proof: unfolding the automaton $\mathcal{A}_{(\sigma,g)}$

$\sigma(a) = ab$
$\sigma(b) = a$

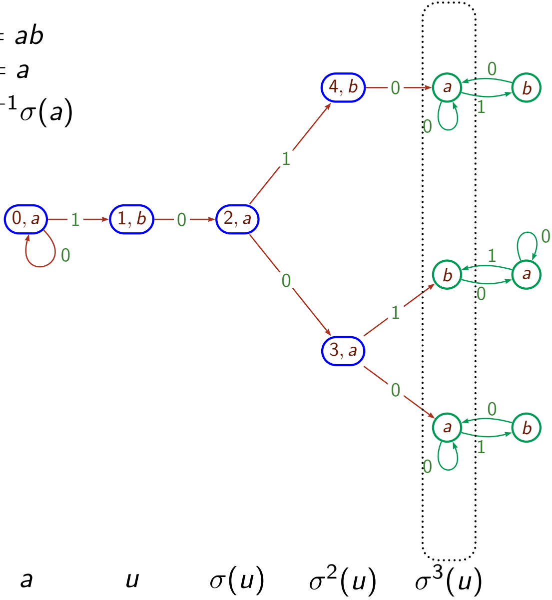
$$u = a^{-1}\sigma(a) \\ = b$$



Idea of proof: unfolding the automaton $\mathcal{A}_{(\sigma,g)}$

$\sigma(a) = ab$
$\sigma(b) = a$

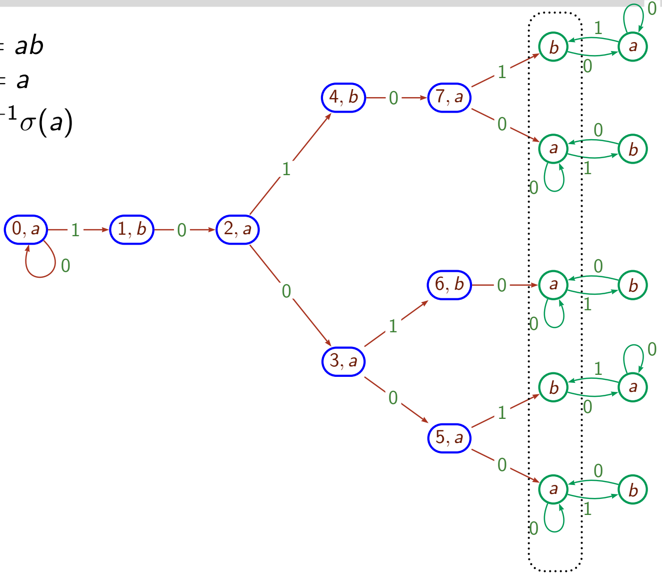
$$u = a^{-1}\sigma(a) = b$$



Idea of proof: unfolding the automaton $\mathcal{A}_{(\sigma,g)}$

$\sigma(a) = ab$
$\sigma(b) = a$

$$u = a^{-1}\sigma(a) = b$$

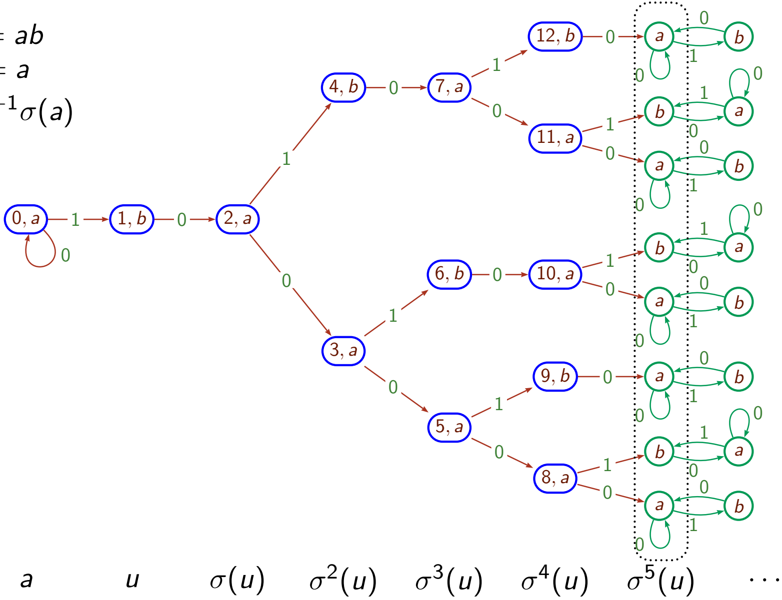


a u $\sigma(u)$ $\sigma^2(u)$ $\sigma^3(u)$ $\sigma^4(u)$

Idea of proof: unfolding the automaton $\mathcal{A}_{(\sigma,g)}$

$\sigma(a) = ab$
$\sigma(b) = a$

$u = a^{-1}\sigma(a)$
$= b$



Backward direction of the theorem

Theorem

L: a prefix-closed language.
Signature(L) is substitutive $\Leftrightarrow$ L is accepted by a finite automaton.

$\mathcal{B}$: a finite automaton.
We define (σ_B, g_B) such that

$$\mathcal{B} \;=\; \mathcal{A}_{(\sigma_B, g_B)}$$

Backward direction of the theorem

Theorem

L: a prefix-closed language.
Signature(L) is substitutive $\Leftrightarrow$ L is accepted by a finite automaton.

$\mathcal{B}$: a finite automaton.
We define (σ_B, g_B) such that

$$\mathcal{B} \;=\; \mathcal{A}_{(\sigma_B, g_B)}$$

Proposition

The language accepted by $\mathcal{B}$ has signature $(\sigma_{\mathcal{B}}, g_{\mathcal{B}})$.

Follows directly from the other direction.

Prefix-closed ARNS

Definition

L: a language over an ordered alphabet A.
The representation $\langle n \rangle_L$ of the integer n in the ANS L is the $(n+1)$-th word of L in the radix order.

In our scheme, $\langle n \rangle_L$ is the word labelling the path $0 \longrightarrow n$.

Prefix-closed ARNS

Definition

L: a Prefix-closed regular language over an ordered alphabet A. The representation $\langle n \rangle_L$ of the integer n in the Prefix-closed ARNS L is the $(n+1)$-th word of L in the radix order.

In our scheme, $\langle n \rangle_L$ is the word labelling the path $0 \longrightarrow n$.

Labelling does not matter

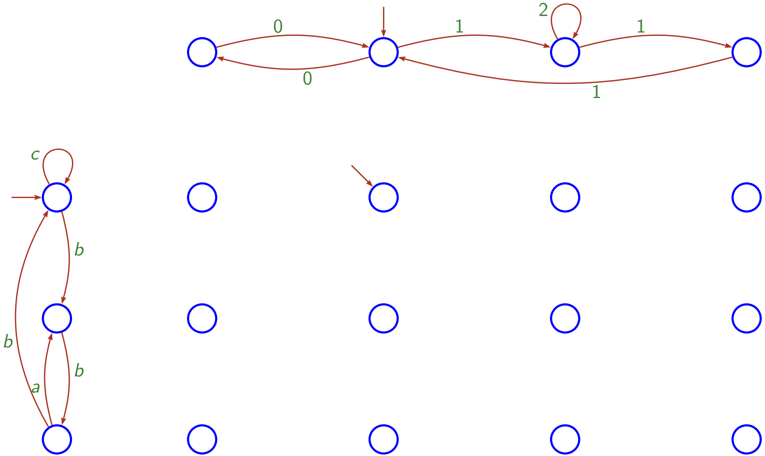
Proposition

L: prefix-closed ARNS of signature (s, λ_1)
K: prefix-closed ARNS of signature (s, λ_2)

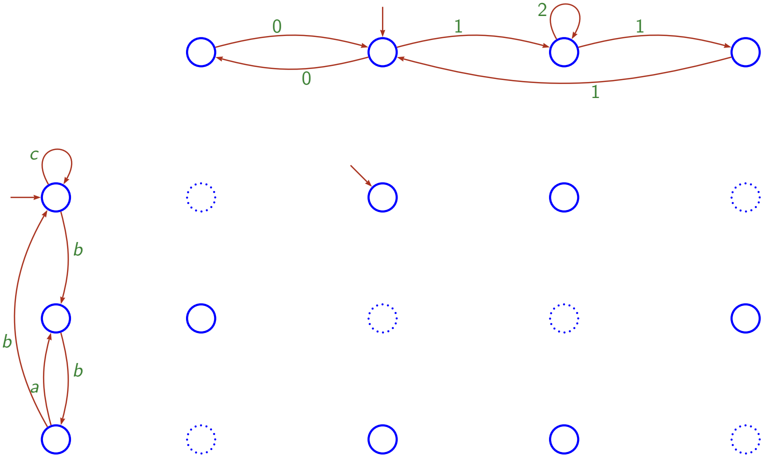
The conversion function $\langle n \rangle_L \mapsto \langle n \rangle_K$ is very simple[†].

[†] realised by a finite, pure sequential and letter-to-letter transducer.

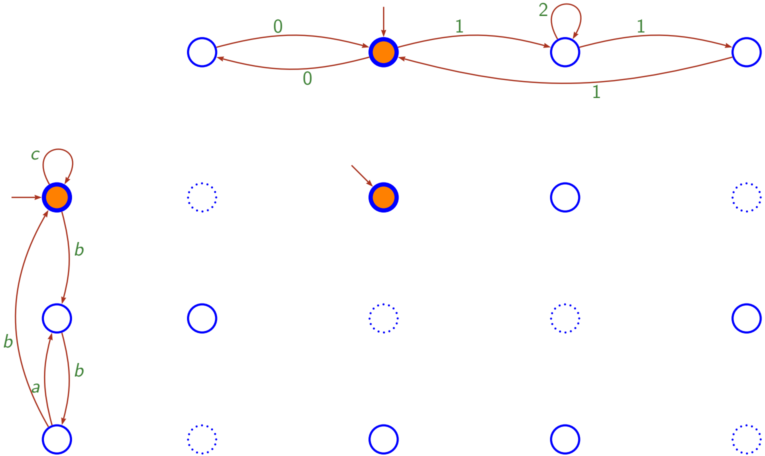
Labelling does not matter (Idea of proof)



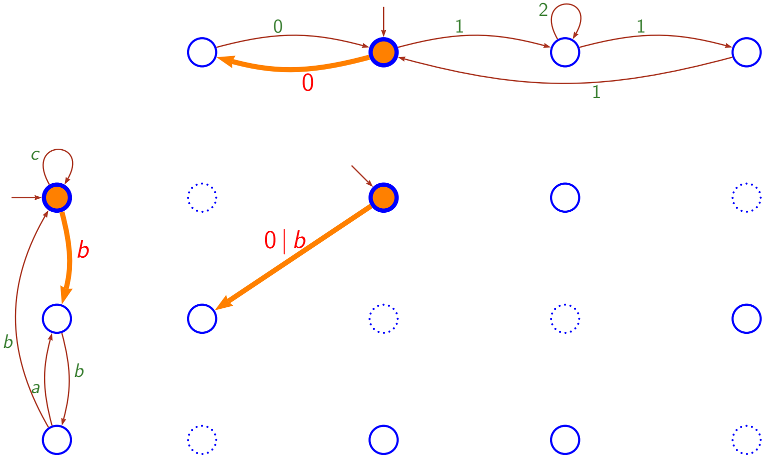
Labelling does not matter (Idea of proof)



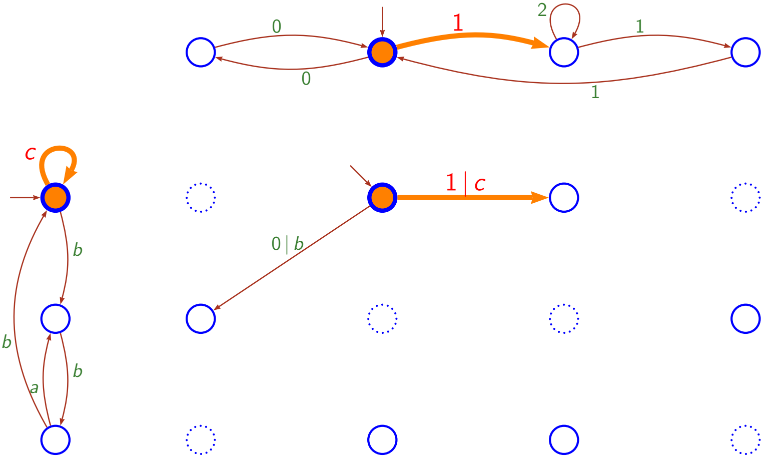
Labelling does not matter (Idea of proof)



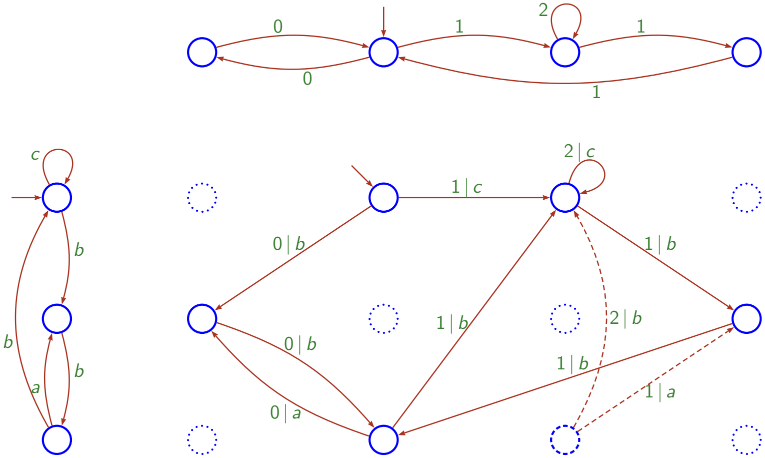
Labelling does not matter (Idea of proof)



Labelling does not matter (Idea of proof)



Labelling does not matter (Idea of proof)



Dumont-Thomas Numeration System (DTNS)

$\sigma : A \to A^*$ prolongable on a.

Example : $\sigma(a) = abc$ $\sigma(b) = ab$ $\sigma(c) = c$

Dumont-Thomas Numeration System (DTNS)

$\sigma : A \to A^*$ prolongable on a.

Example : $\sigma(a) = abc$ $\sigma(b) = ab$ $\sigma(c) = c$

Definition

$$A_\sigma = \{[u] \mid u \text{ is a strict prefix of } \sigma(b) \text{ for some } b \in A\}$$

Example : $A_\sigma = \{ [\varepsilon], [a], [ab] \}$

Dumont-Thomas Numeration System (DTNS)

$\sigma : A \to A^*$ prolongable on a.

Example : $\sigma(a) = abc$ $\sigma(b) = ab$ $\sigma(c) = c$

Definition

$$A_\sigma \;=\; \{[u] \mid u \text{ is a strict prefix of } \sigma(b) \text{ for some } b \in A\}$$

Example : $A_\sigma \;=\; \{\ [\varepsilon], [a], [ab]\ \}$

g_σ: morphism $A^* \to A_\sigma^*$

$g_\sigma(b) = [u_0]\,[u_1] \cdots [u_{k-1}]$

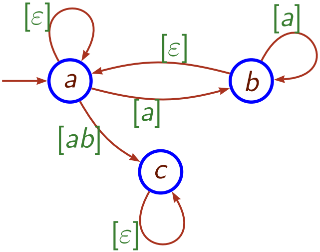
- $k = |\sigma(b)|$
- u_i is the prefix of length i of $\sigma(b)$

Example : $g_\sigma(a) = [\varepsilon]\,[a]\,[ab]$ $g_\sigma(b) = [\varepsilon]\,[a]$ $g_\sigma(c) = [\varepsilon]$

Dumont-Thomas automaton $\mathcal{A}_{(\sigma, g_\sigma)}$

$\sigma(a) = abc$
$\sigma(b) = ab$
$\sigma(c) = c$

$g_\sigma(a) = [\varepsilon]\,[a]\,[ab]$
$g_\sigma(b) = [\varepsilon]\,[a]$
$g_\sigma(c) = [\varepsilon]$



Dumont-Thomas Numeration System (DTNS)

Definition

ρ function ${A_\sigma}^* \to A^*$

$$\rho([u_k]\ \dots\ [u_2]\,[u_1]\,[u_0]) \;=\; \sigma^k(u_k)\sigma^{k-1}(u_{k-1})\cdots\sigma^2(u_2)\sigma(u_1)u_0$$

Example:
$$\begin{array}{lcccl} \rho_\sigma([a]\,[\varepsilon]\,[ab]) = & \sigma^2(a) & \sigma(\varepsilon) & ab & \\ & abc\,abc & \varepsilon & ab & = \text{abc abc ab} \end{array}$$

Dumont-Thomas Numeration System (DTNS)

Definition

ρ function ${A_\sigma}^* \to A^*$

$\rho([u_k]\ \ldots\ [u_2]\,[u_1]\,[u_0]) \;=\; \sigma^k(u_k)\sigma^{k-1}(u_{k-1})\cdots\sigma^2(u_2)\sigma(u_1)u_0$

Example:
$$\begin{array}{lcccccc} \rho_\sigma([a]\,[\varepsilon]\,[ab]) & = & \sigma^2(a) & \sigma(\varepsilon) & ab & & \\ & & abc\,abc & \varepsilon & ab & = & \mathrm{abc\,abc\,ab} \end{array}$$

Theorem (Dumont Thomas ’89)

$\forall n \in \mathbb{N}$

$\exists!$ word $[u_k]\ \ldots\ [u_2]\,[u_1]\,[u_0]$ accepted by $\mathcal{A}_{(\sigma,g_\sigma)}$ such that

- $u_k \neq \varepsilon$
- $|\rho([u_k]\ \ldots\ [u_2]\,[u_1]\,[u_0])| \;=\; n$

$[u_k]\ \ldots\ [u_2]\,[u_1]\,[u_0]$ is the representation of n in the DTNS.

Dumont-Thomas Numeration System (DTNS)

Definition

ρ function ${A_\sigma}^* \to A^*$

$$\rho([u_k]\ \ldots\ [u_2]\,[u_1]\,[u_0]) \ = \ \sigma^k(u_k)\sigma^{k-1}(u_{k-1})\cdots\sigma^2(u_2)\sigma(u_1)u_0$$

Example: $\begin{array}{lllll} \rho_\sigma([a]\,[\varepsilon]\,[ab]) = & \sigma^2(a) & \sigma(\varepsilon) & ab & \\ & abc\,abc & \varepsilon & ab = & \mathrm{abc\,abc\,ab} \end{array}$

Theorem (Dumont Thomas '89)

$\forall n \in \mathbb{N}$

$\exists!$ word $[u_k]\ \ldots\ [u_2]\,[u_1]\,[u_0]$ accepted by $\mathcal{A}_{(\sigma,g_\sigma)}$ such that

- $u_k \neq \varepsilon$
- $|\rho([u_k]\ \ldots\ [u_2]\,[u_1]\,[u_0])| \ = \ n$

$[u_k]\ \ldots\ [u_2]\,[u_1]\,[u_0]$ is the representation of n in the DTNS.

Example: $[a]\,[\varepsilon]\,[ab]$ is the representation of 8 .

Dumont-Thomas Numeration System (DTNS)

Definition

ρ function ${A_\sigma}^* \rightarrow A^*$

$\rho([u_k]\ \ldots\ [u_2]\,[u_1]\,[u_0]) \ = \ \sigma^k(u_k)\sigma^{k-1}(u_{k-1})\cdots\sigma^2(u_2)\sigma(u_1)u_0$

Example: $\rho_\sigma([a]\,[\varepsilon]\,[ab]) = \begin{array}{cccl} \sigma^2(a) & \sigma(\varepsilon) & ab & \\ abc\,abc & \varepsilon & ab & = \mathrm{abc\,abc\,ab} \end{array}$

Theorem (Dumont Thomas '89)

$\forall n \in \mathbb{N}$

$\exists!$ word $[u_k]\ \ldots\ [u_2]\,[u_1]\,[u_0]$ accepted by $\mathcal{A}_{(\sigma,g_\sigma)}$ such that

- $u_k \neq \varepsilon$
- $|\rho([u_k]\ \ldots\ [u_2]\,[u_1]\,[u_0])| \ = \ n$

$[u_k]\ \ldots\ [u_2]\,[u_1]\,[u_0]$ is the representation of n in the DTNS.

Example: $[a]\,[\varepsilon]\,[ab]$ is the representation of 8.

Theorem

1. Every DTNS is a prefix-closed ARNS.
2. Every prefix-closed ARNS is easily[†] convertible to a DTNS.

[†] Through a finite, letter-to-letter and pure sequential transducer.

Theorem

1. Every DTNS is a prefix-closed ARNS.
2. Every prefix-closed ARNS is easily[†] convertible to a DTNS.

[†] Through a finite, letter-to-letter and pure sequential transducer.

Reformulation of 1.

σ: a morphism generating a DTNS.

$$\forall n, p \in \mathbb{N}, \quad \langle n \rangle_\sigma <_{\mathsf{rad}} \langle n+p \rangle_\sigma$$

The proof of 1. is technical and omitted here.

Theorem

1. Every DTNS is a prefix-closed ARNS.
2. Every prefix-closed ARNS is easily[†] convertible to a DTNS.

[†] Through a finite, letter-to-letter and pure sequential transducer.

Reformulation of 1.

σ: a morphism generating a DTNS.

$$\forall n, p \in \mathbb{N}, \quad \langle n \rangle_\sigma <_{\text{rad}} \langle n+p \rangle_\sigma$$

The proof of 1. is technical and omitted here.

Theorem

1. Every DTNS is a prefix-closed ARNS.
2. Every prefix-closed ARNS is easily[†] convertible to a DTNS.

[†] Through a finite, letter-to-letter and pure sequential transducer.

Sketch of proof of 2.

Prefix-Closed ARNS *L*

Theorem

1. Every DTNS is a prefix-closed ARNS.

2. Every prefix-closed ARNS is easily[†] convertible to a DTNS.

[†] Through a finite, letter-to-letter and pure sequential transducer.

Sketch of proof of 2.

$$\begin{tikzcd} \text{Prefix-Closed ARNS } L \arrow[r] & \text{Automaton } \mathcal{A} \end{tikzcd}$$

Theorem

1. Every DTNS is a prefix-closed ARNS.
2. Every prefix-closed ARNS is easily[†] convertible to a DTNS.

[†] Through a finite, letter-to-letter and pure sequential transducer.

Sketch of proof of 2.

$$\begin{tikzcd}
\text{Prefix-Closed ARNS } L \arrow[r] & \text{Automaton } \mathcal{A} \arrow[r] & \text{Morphisms } (\sigma, g)
\end{tikzcd}$$

Theorem

1. Every DTNS is a prefix-closed ARNS.
2. Every prefix-closed ARNS is easily[†] convertible to a DTNS.

[†] Through a finite, letter-to-letter and pure sequential transducer.

Sketch of proof of 2.

Prefix-Closed ARNS L $\longrightarrow$ Automaton $\mathcal{A}$ $\longrightarrow$ Morphisms (σ, g)
of signature (s, λ_1)

where $s = f_\sigma(\sigma^\omega(a))$ $\lambda_1 = g(\sigma^\omega(a))$

Theorem

1. Every DTNS is a prefix-closed ARNS.
2. Every prefix-closed ARNS is easily[†] convertible to a DTNS.

[†] Through a finite, letter-to-letter and pure sequential transducer.

Sketch of proof of 2.

$$\begin{tikzcd}
\begin{matrix}\text{Prefix-Closed ARNS } L \\ \text{of signature } (s, \lambda_1)\end{matrix} \arrow[r] & \text{Automaton } \mathcal{A} \arrow[r] & \text{Morphisms } (\sigma, g) \arrow[d] \\
 & & \text{DTNS } \sigma
\end{tikzcd}$$

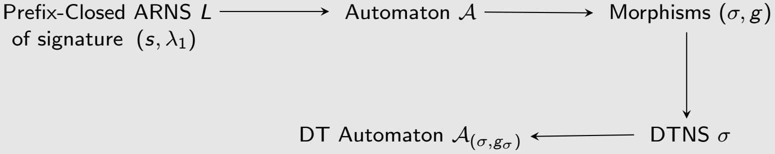
where $s = f_\sigma(\sigma^\omega(a))$ $\lambda_1 = g(\sigma^\omega(a))$

Theorem

1. Every DTNS is a prefix-closed ARNS.
2. Every prefix-closed ARNS is easily[†] convertible to a DTNS.

[†] Through a finite, letter-to-letter and pure sequential transducer.

Sketch of proof of 2.



Prefix-Closed ARNS L of signature (s, λ_1)

Automaton $\mathcal{A}$

Morphisms (σ, g)

DTNS σ

DT Automaton $\mathcal{A}_{(\sigma, g_\sigma)}$

where $s = f_\sigma(\sigma^\omega(a))$ $\lambda_1 = g(\sigma^\omega(a))$

Theorem

1. Every DTNS is a prefix-closed ARNS.
2. Every prefix-closed ARNS is easily[†] convertible to a DTNS.

[†] Through a finite, letter-to-letter and pure sequential transducer.

Sketch of proof of 2.

$$\begin{tikzcd}
\begin{array}{c}\text{Prefix-Closed ARNS } L\\ \text{of signature } (s,\lambda_1)\end{array} \arrow[r] & \text{Automaton } \mathcal{A} \arrow[r] & \text{Morphisms } (\sigma, g) \arrow[d] \\
\begin{array}{c}\text{Prefix-closed ARNS } K\\ \text{of signature } (s,\lambda_2)\end{array} & \text{DT Automaton } \mathcal{A}_{(\sigma, g_\sigma)} \arrow[l] & \text{DTNS } \sigma \arrow[l]
\end{tikzcd}$$

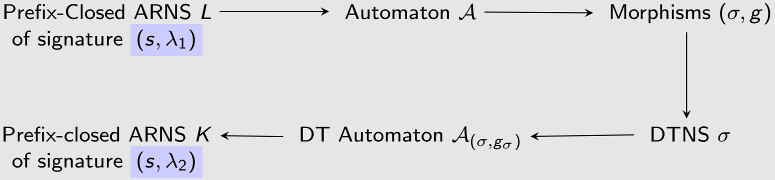
where $s = f_\sigma(\sigma^\omega(a))$ $\lambda_1 = g(\sigma^\omega(a))$ $\lambda_2 = g_\sigma(\sigma^\omega(a))$

Theorem

1. Every DTNS is a prefix-closed ARNS.
2. Every prefix-closed ARNS is easily[†] convertible to a DTNS.

[†] Through a finite, letter-to-letter and pure sequential transducer.

Sketch of proof of 2.



Prefix-Closed ARNS L of signature (s, λ_1)

Automaton $\mathcal{A}$

Morphisms (σ, g)

DTNS σ

DT Automaton $\mathcal{A}_{(\sigma, g_\sigma)}$

Prefix-closed ARNS K of signature (s, λ_2)

where $s = f_\sigma(\sigma^\omega(a))$ $\lambda_1 = g(\sigma^\omega(a))$ $\lambda_2 = g_\sigma(\sigma^\omega(a))$

Proposition

L: prefix-closed ARNS of signature (s, λ_1)
K: prefix-closed ARNS of signature (s, λ_2)
The conversion function $\langle n \rangle_L \mapsto \langle n \rangle_K$ is very simple[†].

Outline

Signature $\mathbf{s} \;=\; (s_0 s_1 \cdots s_{(q-1)})^{\omega}$

- Directing parameter (q, p):
    - the period length of **s** is q;
    - $p \;=\; r_0 + r_1 + r_2 + \cdots + r_{q-1}$.

- Growth ratio: $\frac{p}{q}$
    - Intuition : $\#\{\text{nodes at depth } i\}$ is roughly $\left(\frac{p}{q}\right)^i$

$$\text{Signature } \mathbf{s} \;=\; (s_0 s_1 \cdots s_{(q-1)})^{\omega}$$

- Directing parameter (q, p):
  - the period length of **s** is q;
  - $p \;=\; r_0 + r_1 + r_2 + \cdots + r_{q-1}$.
- Growth ratio: $\frac{p}{q}$
  - Intuition : $\#\{\text{nodes at depth } i\}$ is roughly $\left(\frac{p}{q}\right)^i$

Theorem

$K_{\mathbf{s}}$: the language generated by the signature **s**.

- If $\frac{p}{q}$ is an integer, $K_{\mathbf{s}}$ is a rational language.
  (and linked to integer base NS)
- If $\frac{p}{q}$ is not integer, $K_{\mathbf{s}}$ is a FLIP language.
  (and linked to rational base NS)

Integer Base

- base $p > 1$
- alphabet $A_p = \{0, 1, \cdots, p-1\}$

Integer Base

- base $p > 1$
- alphabet $A_p = \{0, 1, \cdots, p-1\}$

- value $\pi(a_n \cdots a_1 a_0) = \sum_{i=0}^{n} a_i p^i$

Example (base 3) -

$$\begin{aligned} \pi(12) &= (3 \times 1) + (1 \times 2) & = 5 \\ \pi(122) &= (9 \times 1) + (3 \times 2) + (1 \times 2) & = 17 \end{aligned}$$

Integer Base

- base $p > 1$
- alphabet $A_p = \{0, 1, \cdots, p-1\}$

- value $\pi(a_n \cdots a_1 a_0) = \sum_{i=0}^{n} a_i p^i$
- $\pi(A_p^*) = \mathbb{N}$

- representation $\langle n \rangle_p = \langle n' \rangle_p.a$
    - (n', a) is the Euclidean division de n par p.
- $\langle \mathbb{N} \rangle_p = (A_p \backslash \{0\}) \cdot A_p^*$

Rational Base

- base $\frac{p}{q} > 1$ irreducible fraction ($p > q$ and $p \wedge q = 1$).

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
    - (n', a) is the Euclidean division of (**q** $\times n$) by **p** .

Rational Base

- base $\frac{p}{q} > 1$ irreducible fraction ($p > q$ and $p \wedge q = 1$).
- alphabet $A_p = \{0, 1, \ldots, p-1\}$

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of (**q** $\times n$) by **p** .

Rational Base

- base $\frac{p}{q} > 1$ irreducible fraction ($p > q$ and $p \wedge q = 1$).
- alphabet $A_p = \{0, 1, \ldots, p-1\}$

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of (**q** $\times\, n$) by **p** .

Example: computing $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} \quad =$$

Rational Base

- base $\frac{p}{q} > 1$ irreducible fraction ($p > q$ and $p \wedge q = 1$).
- alphabet $A_p = \{0, 1, \ldots, p-1\}$

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
    - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by $\mathbf{p}$.

Example: computing $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} =$$

$$\underset{\substack{\uparrow \\ q}}{2} \times \underset{\substack{\uparrow \\ n}}{3} = \underset{\substack{\uparrow \\ p}}{3} \times N_1 + a_0;$$

Rational Base

- base $\frac{p}{q} > 1$ irreducible fraction ($p > q$ and $p \wedge q = 1$).
- alphabet $A_p = \{0, 1, \ldots, p-1\}$

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of (**q** $\times n$) by **p** .

Example: computing $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} \quad =$$

$$2 \times 3 = 3 \times N_1 + a_0; \qquad \Rightarrow N_1 = 2 \text{ and } a_0 = 0.$$

Rational Base

- base $\frac{p}{q} > 1$ irreducible fraction ($p > q$ and $p \wedge q = 1$).
- alphabet $A_p = \{0, 1, \ldots, p - 1\}$

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of (**q** $\times n$) by **p** .

Example: computing $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} \quad = \quad \langle 2 \rangle_{\frac{3}{2}} 0 \quad =$$

Rational Base

- base $\frac{p}{q} > 1$ irreducible fraction ($p > q$ and $p \wedge q = 1$).
- alphabet $A_p = \{0, 1, \ldots, p-1\}$

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of (**q** $\times n$) by **p** .

Example: computing $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} \quad = \quad \langle 2 \rangle_{\frac{3}{2}} \, 0 \quad =$$

$$2 \times 2 = 3 \times N_2 + a_1;$$

Rational Base

- base $\frac{p}{q} > 1$ irreducible fraction ($p > q$ and $p \wedge q = 1$).
- alphabet $A_p = \{0, 1, \dots, p-1\}$

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of (**q** $\times n$) by **p** .

Example: computing $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} \quad = \quad \langle 2 \rangle_{\frac{3}{2}} 0 \quad =$$

$2 \times 2 = 3 \times N_2 + a_1; \qquad \Rightarrow N_2 = 1$ and $a_1 = 1$.

Rational Base

- base $\frac{p}{q} > 1$ irreducible fraction ($p > q$ and $p \wedge q = 1$).
- alphabet $A_p = \{0, 1, \ldots, p-1\}$

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of (**q** $\times n$) by **p** .

Example: computing $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} \quad = \quad \langle 2 \rangle_{\frac{3}{2}} 0 \quad = \quad \langle 1 \rangle_{\frac{3}{2}} 10 \quad =$$

Rational Base

- base $\frac{p}{q} > 1$ irreducible fraction ($p > q$ and $p \wedge q = 1$).
- alphabet $A_p = \{0, 1, \ldots, p-1\}$

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}} . a$:
  - (n', a) is the Euclidean division of (**q** $\times\, n$) by **p** .

Example: computing $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} \quad = \quad \langle 2 \rangle_{\frac{3}{2}}\, 0 \quad = \quad \langle 1 \rangle_{\frac{3}{2}}\, 10 \quad =$$

$2 \times 1 = 3 \times N_3 + a_2$;

Rational Base

- base $\frac{p}{q} > 1$ irreducible fraction ($p > q$ and $p \wedge q = 1$).
- alphabet $A_p = \{0, 1, \ldots, p-1\}$

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of (**q** $\times n$) by **p** .

Example: computing $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} \quad = \quad \langle 2 \rangle_{\frac{3}{2}} 0 \quad = \quad \langle 1 \rangle_{\frac{3}{2}} 10 \quad =$$

$2 \times 1 = 3 \times N_3 + a_2; \qquad \Rightarrow N_3 = 0$ and $a_2 = 2$.

Rational Base

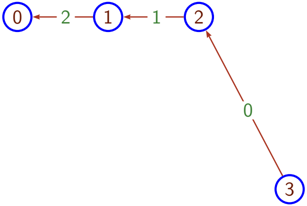
- base $\frac{p}{q} > 1$ irreducible fraction ($p > q$ and $p \wedge q = 1$).
- alphabet $A_p = \{0, 1, \ldots, p-1\}$

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}} .a$:
  - (n', a) is the Euclidean division of (**q** $\times n$) by **p** .

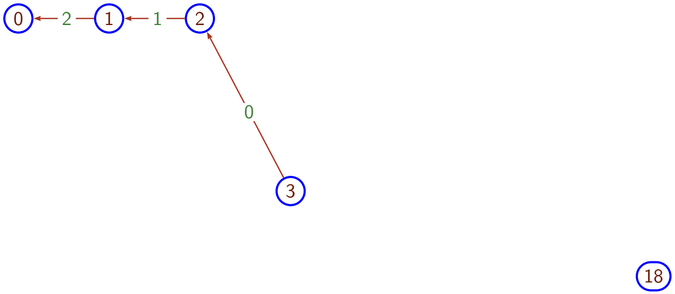
Example: computing $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} \quad = \quad \langle 2 \rangle_{\frac{3}{2}} 0 \quad = \quad \langle 1 \rangle_{\frac{3}{2}} 10 \quad = \quad 210$$

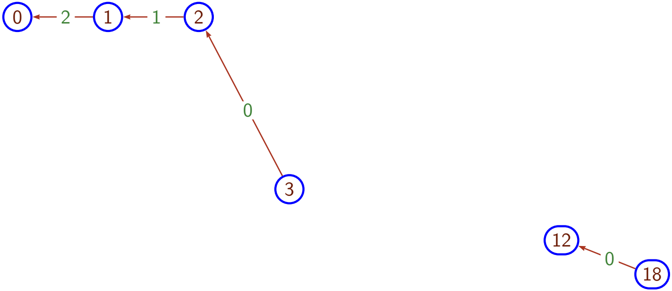
Tree of the representations in base $\frac{3}{2}$



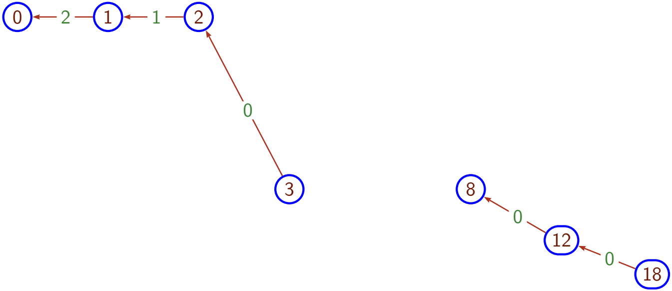
Tree of the representations in base $\frac{3}{2}$



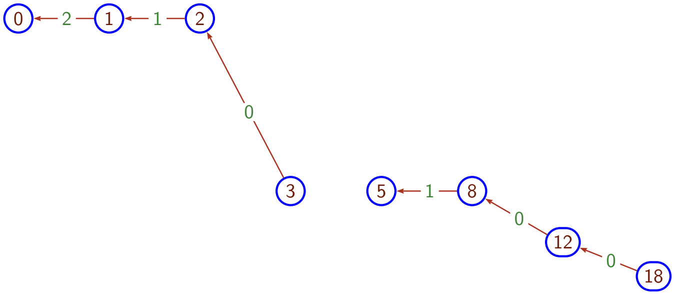
Tree of the representations in base $\frac{3}{2}$



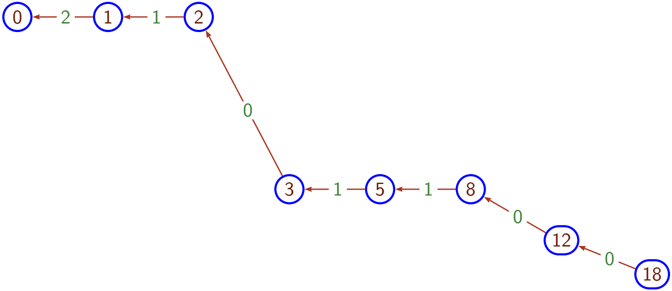
Tree of the representations in base $\frac{3}{2}$



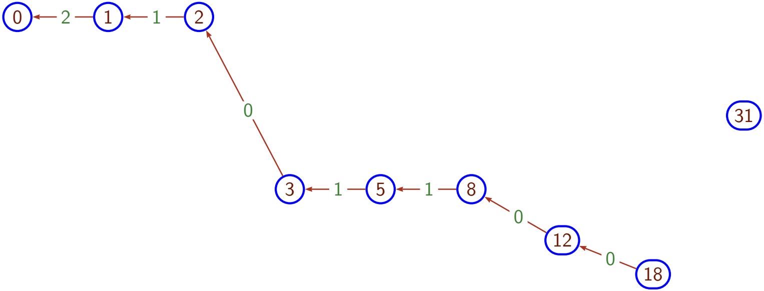
Tree of the representations in base $\frac{3}{2}$



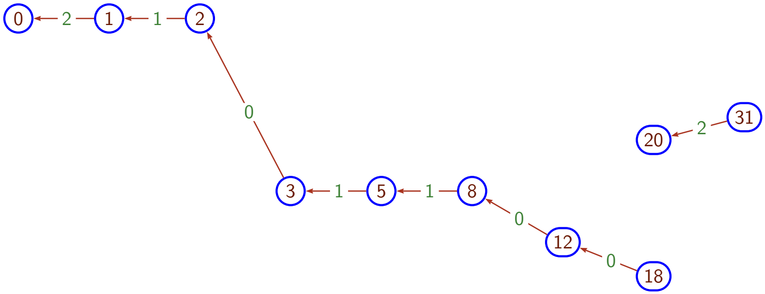
Tree of the representations in base $\frac{3}{2}$



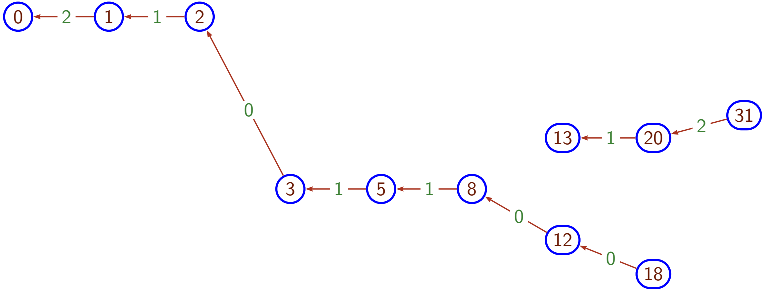
Tree of the representations in base $\frac{3}{2}$



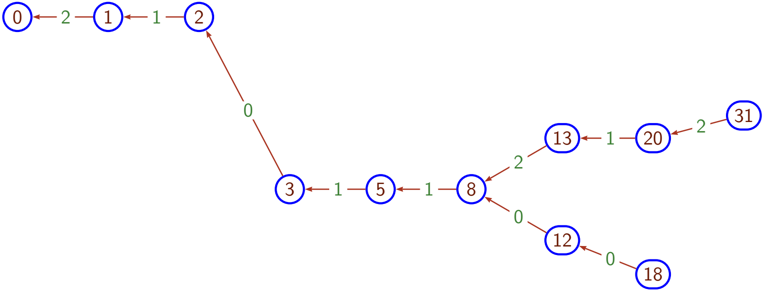
Tree of the representations in base $\frac{3}{2}$



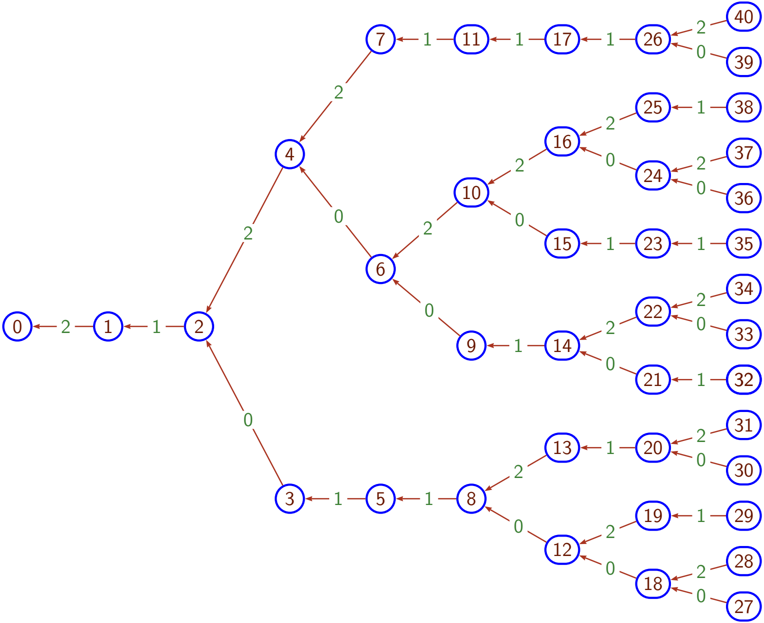
Tree of the representations in base $\frac{3}{2}$



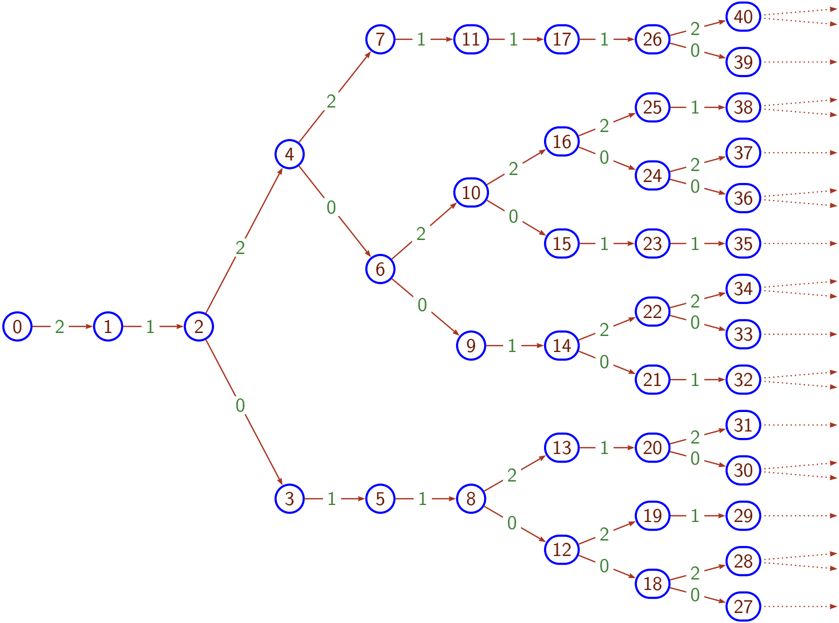
Tree of the representations in base $\frac{3}{2}$



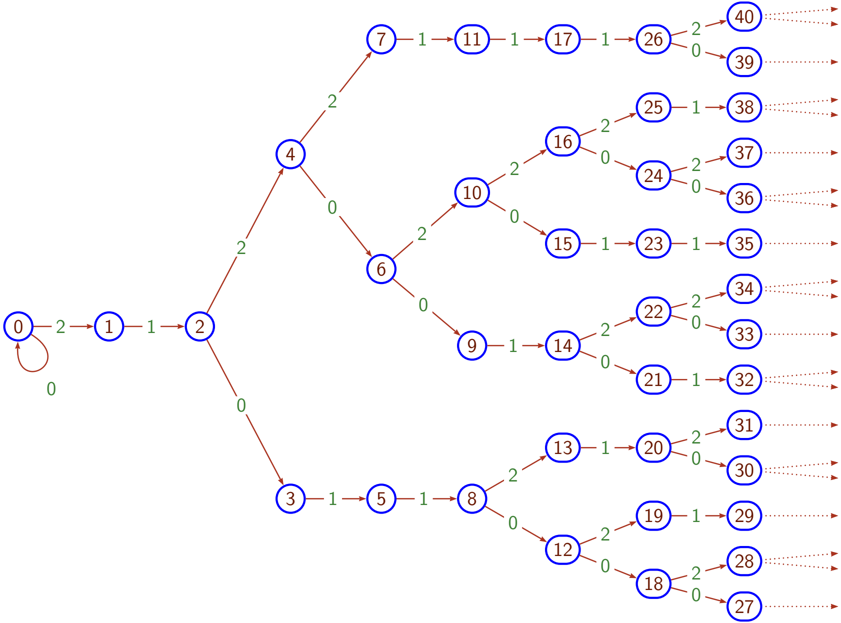
Tree of the representations in base $\frac{3}{2}$



The language $L_{\frac{3}{2}}$



The language $L_{\frac{3}{2}}$



Properties of $L_{\frac{p}{q}}$

- $L_{\frac{p}{q}}$ is prefix-closed.
- $L_{\frac{p}{q}}$ is right-extendable.

Properties of $L_{\frac{p}{q}}$

- $L_{\frac{p}{q}}$ is prefix-closed.
- $L_{\frac{p}{q}}$ is right-extendable.

Let $a_n a_{n-1} \cdots a_0 = \langle n \rangle$.

$$\sum_{i=0}^{n} \frac{a_i}{q} \left(\frac{p}{q}\right)^i \quad = \quad n$$

Properties of $L_{\frac{p}{q}}$

- $L_{\frac{p}{q}}$ is prefix-closed.
- $L_{\frac{p}{q}}$ is right-extendable.

Let $a_n a_{n-1} \cdots a_0 = \langle n \rangle$.

$$\pi(a_n \cdots a_1 a_0) \;=\; \textstyle\sum_{i=0}^{n} \frac{a_i}{q} \left(\frac{p}{q}\right)^i \;=\; n$$

Properties of $L_{\frac{p}{q}}$

- $L_{\frac{p}{q}}$ is prefix-closed.
- $L_{\frac{p}{q}}$ is right-extendable.

Let $a_n a_{n-1} \cdots a_0 = \langle n \rangle$.

$$\pi(a_n \cdots a_1 a_0) \;=\; \sum_{i=0}^{n} \frac{a_i}{q} \left(\frac{p}{q}\right)^i \;=\; n$$

Theorem (Akiyama Frougny Sakarovitch, 2008)

$L_{\frac{p}{q}}$ is not a context-free language.

Properties of $L_{\frac{p}{q}}$

- $L_{\frac{p}{q}}$ is prefix-closed.
- $L_{\frac{p}{q}}$ is right-extendable.

Let $a_n a_{n-1} \cdots a_0 = \langle n \rangle$.

$$\pi(a_n \cdots a_1 a_0) \;=\; \textstyle\sum_{i=0}^{n} \frac{a_i}{q} \left(\frac{p}{q}\right)^i \;=\; n$$

Theorem (Akiyama Frougny Sakarovitch, 2008)

$L_{\frac{p}{q}}$ is not a context-free language.

$L_{\frac{p}{q}}$ has the Finite Left Iteration Property.

The Finite Left Iteration Property (FLIP)

Definition

A language L is FLIP if

$\forall u\ v,\ \ \exists$ only finitely indices i

such that $u\,v^i$ is the prefix of a word of L;

or, equivalently

$\forall u\ v,\quad \mathrm{Pref}(L) \bigcap u\,v^* \quad$ is finite

Example : the prefixes of an infinite aperiodic word.

(We are still looking for “natural” examples of FLIP languages.)

The Finite Left Iteration Property (FLIP)

Definition

A language L is FLIP if
$\forall u\ v$, $\exists$ only finitely indices i
such that $u\,v^i$ is the prefix of a word of L;

or, equivalently
$\forall u\ v$, $\mathrm{Pref}(L) \bigcap u\,v^*$ is finite

Intuition 1

- L does not contain any infinite rational language.
  [IRS : Greibach 1975]
- L is "hard" to extend to a rational language.

Example: $\{a^n \mid n \text{ is a prime number}\}$ is IRS but not FLIP.

The Finite Left Iteration Property (FLIP)

Definition

A language L is FLIP if

$$\forall u\ v, \quad \exists \text{ only finitely indices } i$$
$$\text{such that } u\,v^i \text{ is the prefix of a word of } L;$$

or, equivalently

$$\forall u\ v, \quad \mathrm{Pref}(L) \bigcap u\,v^* \quad \text{is finite}$$

Intuition 2

- The topological closure of L contains **only** aperiodic word.

(Every branch of the tree-representation of L is labelled by an aperiodic word.)

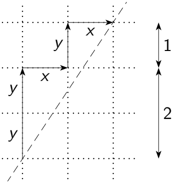
FLIP is a very robust property

- Every finite language is FLIP.

- A **finite union** of FLIP languages is FLIP.
- Any **intersection** of FLIP languages is FLIP.
- Every **sub-language** of a FLIP language is FLIP.
- The **concatenation** of two FLIP languages is FLIP.

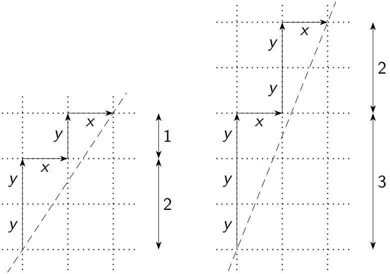
- The **prefix closure** of a FLIP language is FLIP.
- The **inverse image by transducer** of a FLIP language is FLIP.

Christoffel Word and Christoffel rhythm



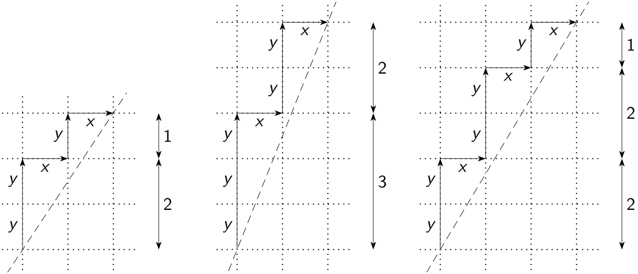
Slope $\frac{p}{q}$: $\frac{3}{2}$
Christ. word: $yy\,x\,y\,x$
Christ. rhythm: (2, 1)
Sign. of $L_{\frac{p}{q}}$: $(21)^{\omega}$

Christoffel Word and Christoffel rhythm



Slope $\frac{p}{q}$:	$\frac{3}{2}$	$\frac{5}{2}$
Christ. word:	$yy\ x\ y\ x$	$yyy\ x\ yy\ x$
Christ. rhythm:	(2, 1)	(3, 2)
Sign. of $L_{\frac{p}{q}}$:	$(21)^{\omega}$	$(32)^{\omega}$

Christoffel Word and Christoffel rhythm



Slope $\frac{p}{q}$:	$\frac{3}{2}$	$\frac{5}{2}$	$\frac{5}{3}$
Christ. word:	$yy\ x\ y\ x$	$yyy\ x\ yy\ x$	$yy\ x\ yy\ x\ y\ x$
Christ. rhythm:	(2, 1)	(3, 2)	(2, 2, 1)
Sign. of $L_{\frac{p}{q}}$:	$(21)^{\omega}$	$(32)^{\omega}$	$(221)^{\omega}$

Definition (Canonical labelling)

the p-tuple: $(0, q, (2q), \ldots, (p-1)q) \quad [mod \ p]$

Example: $(0, 2, 1)$ for $\frac{3}{2}$ and $(0, 3, 1, 4, 2)$ for $\frac{5}{3}$.

Definition (Canonical labelling)

the p-tuple: $(0, q, (2q), \ldots, (p-1)q) \quad [mod\ p]$

Example: $(0, 2, 1)$ for $\frac{3}{2}$ and $(0, 3, 1, 4, 2)$ for $\frac{5}{3}$.

Proposition (MS, to appear)

$\frac{p}{q}$: a base.
u: the Christoffel rhythm of slope $\frac{p}{q}$.
v: the canonical labelling associated with $\frac{p}{q}$.
The language $L_{\frac{p}{q}}$ has for signature u^ω and for labelling v^ω.

The proof is technical and omitted here.

$$\text{Signature } \mathbf{s} = (s_0 s_1 \cdots s_{(q-1)})^\omega$$

- Directing parameter (q, p):
  - the period length of **s** is q;
  - $p = s_0 + s_1 + s_2 + \cdots + s_{q-1}$.

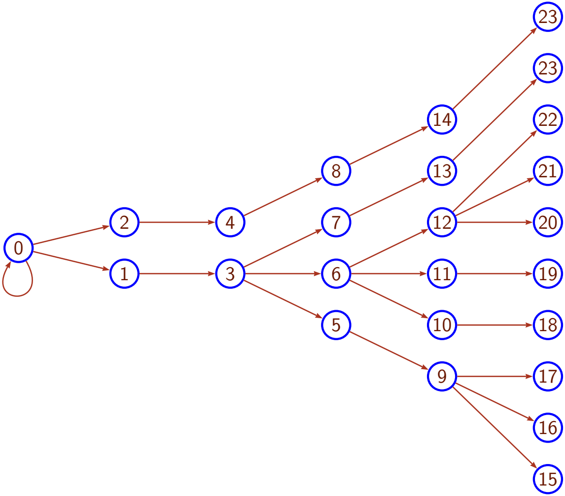
- Growth ratio: $\frac{p}{q}$
  - Intuition : $\#\{\text{nodes at depth } i\}$ is roughly $\left(\frac{p}{q}\right)^i$

Theorem (MS, to appear)

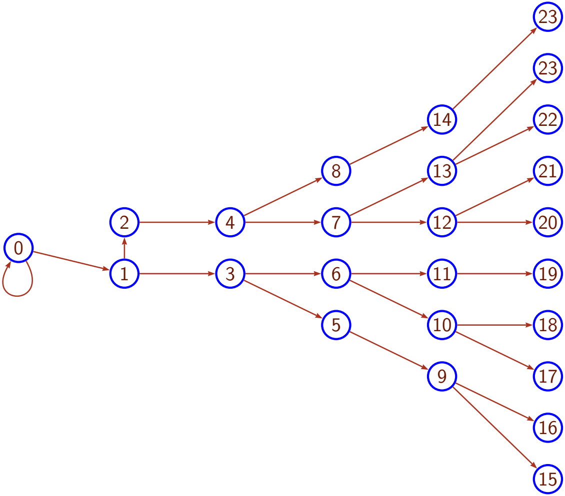
$K_\mathbf{s}$: the language generated by the signature **s**.

- If $\frac{p}{q}$ is an integer, $K_\mathbf{s}$ is a rational language.
  (and linked to integer base NS)
- If $\frac{p}{q}$ is not integer, $K_\mathbf{s}$ is a FLIP language.
  (and linked to rational base NS)

The tree whose signature is $(3, 1, 1)^\omega$



The tree whose signature is $(2, 2, 1)^\omega$



The tree whose signature is $(2, 2, 1)^{\omega}$

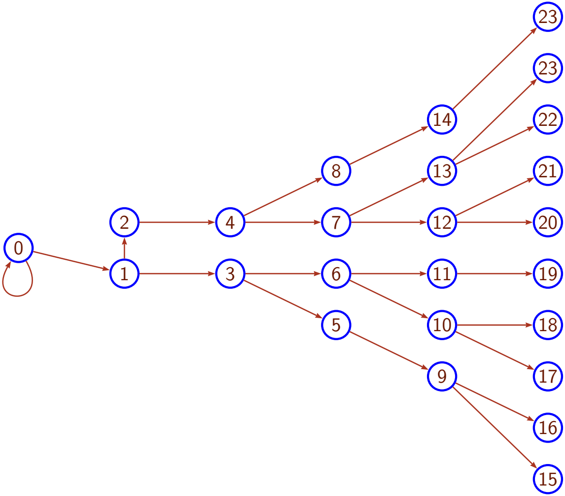


Figure: Underlying tree of the language of integers in base $\frac{5}{3}$

Signature $\mathbf{s} = (s_0 s_1 \cdots s_{(q-1)})^\omega$

- Directing parameter (q, p):
  - the period length of **s** is q;
  - $p = s_0 + s_1 + s_2 + \cdots + s_{q-1}$.

- Growth ratio: $\frac{p}{q}$
  - Intuition : $\#\{\text{nodes at depth } i\}$ is roughly $\left(\frac{p}{q}\right)^i$

Signature $\mathbf{s} = (s_0 s_1 \cdots s_{(q-1)})^{\omega}$

- Directing parameter (q, p):
  - the period length of **s** is q;
  - $p = s_0 + s_1 + s_2 + \cdots + s_{q-1}$.

- Growth ratio: $\frac{p}{q}$
  - Intuition : $\#\{\text{nodes at depth } i\}$ is roughly $\left(\frac{p}{q}\right)^i$
- Smart labelling: $(\lambda_0 \lambda_1 \cdots \lambda_{p-1})^{\omega}$

  $\lambda_0 = 0$

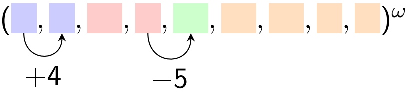
  Inside a block : $\lambda_{i+1} = \lambda_i + q$

  From a block to the next : $\lambda_{i+1} = \lambda_i + q - p$.

  (see example)

Example: $\mathbf{s} = (2, 2, 1, 4)^{\omega}$ whose dir. par. is $(4, 9)$.

Smart labelling :



Signature $\mathbf{s} \;=\; (s_0 s_1 \cdots s_{(q-1)})^{\omega}$

- Directing parameter (q, p):
  - the period length of $\mathbf{s}$ is q;
  - $p \;=\; s_0 + s_1 + s_2 + \cdots + s_{q-1}$.
- Growth ratio: $\frac{p}{q}$
  - Intuition : $\#\{\text{nodes at depth } i\}$ is roughly $\left(\frac{p}{q}\right)^i$
- Smart labelling: $(\lambda_0 \lambda_1 \cdots \lambda_{p-1})^{\omega}$

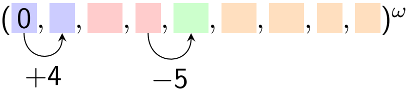
  $\lambda_0 = 0$

  Inside a block : $\lambda_{i+1} = \lambda_i + q$

  From a block to the next : $\lambda_{i+1} = \lambda_i + q - p$.

  (see example)

Example: $\mathbf{s} = (\boxed{2}, \boxed{2}, \boxed{1}, \boxed{4})^{\omega}$ whose dir. par. is $(4, 9)$.


Smart labelling : $(\boxed{0}, \boxed{}, \boxed{}, \boxed{}, \boxed{}, \boxed{}, \boxed{}, \boxed{}, \boxed{})^{\omega}$

Signature $\mathbf{s} \; = \; (s_0 s_1 \cdots s_{(q-1)})^{\omega}$

- Directing parameter (q, p):
  - the period length of **s** is q;
  - $p \; = \; s_0 + s_1 + s_2 + \cdots + s_{q-1}$.

- Growth ratio: $\frac{p}{q}$
  - Intuition : $\#\{$nodes at depth $i\}$ is roughly $\left(\frac{p}{q}\right)^{i}$

- Smart labelling: $(\lambda_0 \lambda_1 \cdots \lambda_{p-1})^{\omega}$

  $\lambda_0 = 0$

  Inside a block : $\lambda_{i+1} = \lambda_i + q$

  From a block to the next : $\lambda_{i+1} = \lambda_i + q - p$.

  (see example)

Example: $\mathbf{s} = (\boxed{2}, \boxed{2}, \boxed{1}, \boxed{4})^{\omega}$ whose dir. par. is $(4, 9)$.

Smart labelling : $(\boxed{0}, \boxed{4}, \boxed{}, \boxed{}, \boxed{}, \boxed{}, \boxed{}, \boxed{}, \boxed{})^{\omega}$

$+4$ -5

Signature $\mathbf{s} = (s_0 s_1 \cdots s_{(q-1)})^\omega$

- Directing parameter (q, p):
  - the period length of **s** is q;
  - $p = s_0 + s_1 + s_2 + \cdots + s_{q-1}$.

- Growth ratio: $\frac{p}{q}$
  - Intuition : $\#\{$nodes at depth $i\}$ is roughly $\left(\frac{p}{q}\right)^i$

- Smart labelling: $(\lambda_0 \lambda_1 \cdots \lambda_{p-1})^\omega$

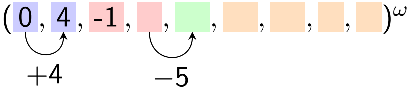
  $\lambda_0 = 0$

  Inside a block : $\lambda_{i+1} = \lambda_i + q$

  From a block to the next : $\lambda_{i+1} = \lambda_i + q - p$.

  (see example)

Example: $\mathbf{s} = (2, 2, 1, 4)^\omega$ whose dir. par. is $(4, 9)$.



Smart labelling : $(0, 4, \text{-1},$ 🟥, 🟩, 🟧, 🟧, 🟧, 🟧$)^\omega$

Signature $\mathbf{s} = (s_0 s_1 \cdots s_{(q-1)})^\omega$

- Directing parameter (q, p):
  - the period length of **s** is q;
  - $p = s_0 + s_1 + s_2 + \cdots + s_{q-1}$.
- Growth ratio: $\frac{p}{q}$
  - Intuition : $\#\{$nodes at depth $i\}$ is roughly $\left(\frac{p}{q}\right)^i$
- Smart labelling: $(\lambda_0 \lambda_1 \cdots \lambda_{p-1})^\omega$

  $\lambda_0 = 0$

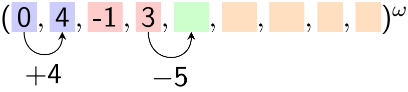
  Inside a block : $\lambda_{i+1} = \lambda_i + q$

  From a block to the next : $\lambda_{i+1} = \lambda_i + q - p$.

  (see example)

Example: $\mathbf{s} = (2, 2, 1, 4)^\omega$ whose dir. par. is $(4, 9)$.

Smart labelling : $(0, 4, -1, 3, \ , \ , \ , \ , \ )^\omega$



Signature $\mathbf{s} = (s_0 s_1 \cdots s_{(q-1)})^\omega$

- Directing parameter (q, p):
  - the period length of **s** is q;
  - $p = s_0 + s_1 + s_2 + \cdots + s_{q-1}$.

- Growth ratio: $\frac{p}{q}$
  - Intuition : $\#\{\text{nodes at depth } i\}$ is roughly $\left(\frac{p}{q}\right)^i$

- Smart labelling: $(\lambda_0 \lambda_1 \cdots \lambda_{p-1})^\omega$
  $\lambda_0 = 0$
  Inside a block : $\lambda_{i+1} = \lambda_i + q$
  From a block to the next : $\lambda_{i+1} = \lambda_i + q - p$.
  (see example)

Example: $\mathbf{s} = (2, 2, 1, 4)^\omega$ whose dir. par. is $(4, 9)$.

Smart labelling : $(\underset{+4}{\underbrace{0, 4}}, -1, \underset{-5}{\underbrace{3, -2}}, \ , \ , \ , \ )^\omega$

Signature $\mathbf{s} \; = \; (s_0 s_1 \cdots s_{(q-1)})^\omega$

- Directing parameter (q, p):
    - the period length of **s** is q;
    - $p \; = \; s_0 + s_1 + s_2 + \cdots + s_{q-1}$.
- Growth ratio: $\frac{p}{q}$
    - Intuition : $\#\{$nodes at depth $i\}$ is roughly $\left(\frac{p}{q}\right)^i$
- Smart labelling: $(\lambda_0 \lambda_1 \cdots \lambda_{p-1})^\omega$
  $\lambda_0 = 0$
  Inside a block : $\lambda_{i+1} = \lambda_i + q$
  From a block to the next : $\lambda_{i+1} = \lambda_i + q - p$.
  (see example)

Example: $\mathbf{s} = (2, 2, 1, 4)^\omega$ whose dir. par. is $(4, 9)$.

Smart labelling : $(0, 4, \text{-}1, 3, \text{-}2, \text{-}7, , ,)^\omega$

$+4$ $\quad$ -5

Signature $\mathbf{s} \;=\; (s_0 s_1 \cdots s_{(q-1)})^{\omega}$

- Directing parameter (q, p):
  - the period length of **s** is q;
  - $p \;=\; s_0 + s_1 + s_2 + \cdots + s_{q-1}$.

- Growth ratio: $\frac{p}{q}$
  - Intuition : $\#\{\text{nodes at depth } i\}$ is roughly $\left(\frac{p}{q}\right)^{i}$

- Smart labelling: $(\lambda_0 \lambda_1 \cdots \lambda_{p-1})^{\omega}$

  $\lambda_0 = 0$

  Inside a block : $\lambda_{i+1} = \lambda_i + q$

  From a block to the next : $\lambda_{i+1} = \lambda_i + q - p$.

  (see example)

Example: $\mathbf{s} = (2, 2, 1, 4)^{\omega}$ whose dir. par. is $(4, 9)$.

Smart labelling : $(0, 4, \text{-}1, 3, \text{-}2, \text{-}7, \text{-}3, \;, \;)^{\omega}$

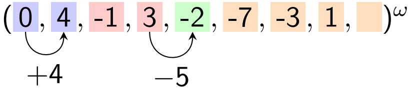
$+4$ -5

Signature $\mathbf{s} = (s_0 s_1 \cdots s_{(q-1)})^\omega$

- Directing parameter (q,p):
  - the period length of **s** is q;
  - $p = s_0 + s_1 + s_2 + \cdots + s_{q-1}$.

- Growth ratio: $\frac{p}{q}$
  - Intuition : $\#\{\text{nodes at depth } i\}$ is roughly $\left(\frac{p}{q}\right)^i$
- Smart labelling: $(\lambda_0 \lambda_1 \cdots \lambda_{p-1})^\omega$
  $\lambda_0 = 0$
  Inside a block : $\lambda_{i+1} = \lambda_i + q$
  From a block to the next : $\lambda_{i+1} = \lambda_i + q - p$.
  (see example)

Example: $\mathbf{s} = (2, 2, 1, 4)^\omega$ whose dir. par. is $(4,9)$.



Smart labelling : $(0, 4, \text{-}1, 3, \text{-}2, \text{-}7, \text{-}3, 1, \ )^\omega$

Signature $\mathbf{s} = (s_0 s_1 \cdots s_{(q-1)})^{\omega}$

- Directing parameter (q, p):
  - the period length of **s** is q;
  - $p = s_0 + s_1 + s_2 + \cdots + s_{q-1}$.

- Growth ratio: $\frac{p}{q}$
  - Intuition : $\#\{\text{nodes at depth } i\}$ is roughly $\left(\frac{p}{q}\right)^i$

- Smart labelling: $(\lambda_0 \lambda_1 \cdots \lambda_{p-1})^{\omega}$
  $\lambda_0 = 0$
  Inside a block : $\lambda_{i+1} = \lambda_i + q$
  From a block to the next : $\lambda_{i+1} = \lambda_i + q - p$.
  (see example)

Example: $\mathbf{s} = (2, 2, 1, 4)^{\omega}$ whose dir. par. is $(4, 9)$.

Smart labelling : $(0, 4, -1, 3, -2, -7, -3, 1, 5)^{\omega}$

$+4$ -5

Proposition (MS, to appear)

u^ω: a periodic signature.
(q, p): its directing parameter.
v^ω: its associated smart labelling
L: the language whose signature/labelling are (u^ω, v^ω)

if w is the $(n+1)$-th word of L (labelling the path $0 \xrightarrow{w} n$)

$$\pi_{\frac{p}{q}}(w) = n$$

L is a "non-canonical representation of integers" in base $\frac{p}{q}$

Reformulation of Theorem 2.

L: a non canonical representation of integers in base $\frac{p}{q}$.
L is FLIP.

Theorem (Akiyama Frougny Sakarovitch 2008)

For all finite alphabet A there is a finite sequential transducer $\mathcal{T}$:
$\forall w \in A^*, \qquad \pi(w) = \pi(\mathcal{T}(w)) \qquad \text{and} \qquad \mathcal{T}(w) \in L_{\frac{p}{q}}.$

- It follows that $\mathcal{T}(L) = L_{\frac{p}{q}}$
- FLIP is stable by inverse image of transducer
  hence $\mathcal{T}^{-1}(L_{\frac{p}{q}})$ is FLIP.
- FLIP is stable by sublanguage
  hence L is FLIP.

Outline

Ultimately periodic signature

Periodic Signature

Example: $\mathbf{s} = (2\ 2\ 1\ 4)^{\omega}$ whose dir. par. is $(4, 9)$.

Smart labelling :

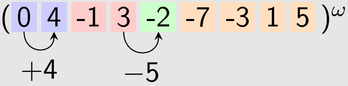
$$\begin{tikzcd}[column sep=small] (& 0 \arrow[r, bend right, "+4"'] & 4 & -1 & 3 \arrow[r, bend right, "-5"'] & -2 & -7 & -3 & 1 & 5 &)^{\omega} \end{tikzcd}$$

Ultimately periodic signature

Periodic Signature

Example: $\mathbf{s} = (2\ 2\ 1\ 4)^{\omega}$ whose dir. par. is $(4, 9)$.

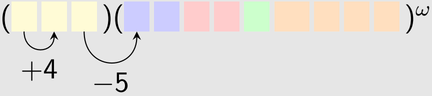
Smart labelling : $(0\ 4\ \text{-1}\ 3\ \text{-2}\ \text{-7}\ \text{-3}\ 1\ 5)^{\omega}$



Ultimately Periodic Signature

Example: $\mathbf{s} = 3(2\ 2\ 1\ 4)^{\omega}$ whose dir. par. is $(4, 9)$.

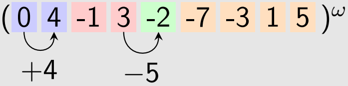
Smart labelling :



Ultimately periodic signature

Periodic Signature

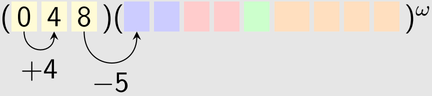
Example: $\mathbf{s} = (2\ 2\ 1\ 4)^{\omega}$ whose dir. par. is $(4, 9)$.



Smart labelling : $(0\ 4\ \text{-}1\ 3\ \text{-}2\ \text{-}7\ \text{-}3\ 1\ 5)^{\omega}$

Ultimately Periodic Signature

Example: $\mathbf{s} = 3(2\ 2\ 1\ 4)^{\omega}$ whose dir. par. is $(4, 9)$.



Smart labelling : $(0\ 4\ 8)($🟦🟦🟥🟥🟩🟧🟧🟧🟧$)^{\omega}$

Ultimately periodic signature

Periodic Signature

Example: $\mathbf{s} = (2\ 2\ 1\ 4)^\omega$ whose dir. par. is $(4, 9)$.

Smart labelling : $(0\ 4\ \text{-}1\ 3\ \text{-}2\ \text{-}7\ \text{-}3\ 1\ 5\ )^\omega$

$+4$ -5

Ultimately Periodic Signature

Example: $\mathbf{s} = 3(2\ 2\ 1\ 4)^\omega$ whose dir. par. is $(4, 9)$.

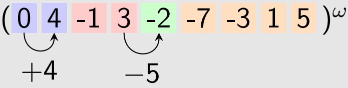
Smart labelling : $(0\ 4\ 8\ )(3$ 🟦 🟥 🟥 🟩 🟧 🟧 🟧 🟧 $)^\omega$

$+4$ -5

Ultimately periodic signature

Periodic Signature

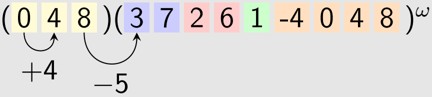
Example: $\mathbf{s} = (2\ 2\ 1\ 4)^\omega$ whose dir. par. is $(4, 9)$.



Smart labelling : $(0\ 4\ \text{-}1\ 3\ \text{-}2\ \text{-}7\ \text{-}3\ 1\ 5\ )^\omega$

Ultimately Periodic Signature

Example: $\mathbf{s} = 3(2\ 2\ 1\ 4)^\omega$ whose dir. par. is $(4, 9)$.

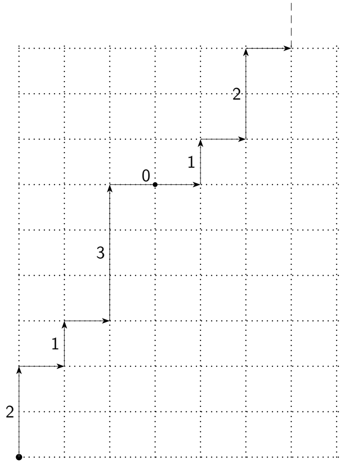


Smart labelling : $(0\ 4\ 8\ )(3\ 7\ 2\ 6\ 1\ \text{-}4\ 0\ 4\ 8\ )^\omega$

→ Will also generate a “non-canonical representation of integers” in base $\frac{9}{4}$, hence a FLIP language.

Going to the limit: signature directed by $\frac{p}{q}$

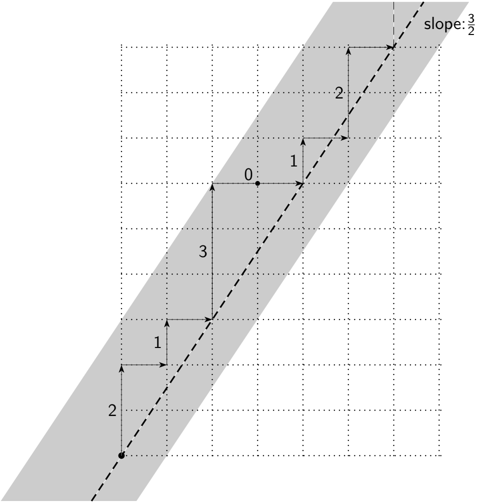
$\mathbf{s} = 2\,1\,3\,0\,1\,2\cdots$



Going to the limit: signature directed by $\frac{p}{q}$

$\mathbf{s} = 2\,1\,3\,0\,1\,2\cdots$

$\mathbf{s}$ is directed by $\frac{3}{2}$



s directed by β

- β belongs to $\mathbb{Q} \setminus \mathbb{N}$
    - linked to rational base number system;
    - non-canonical representation;
    - always a FLIP Language.
- β belongs to $\mathbb{N}$
    - linked to integer base *b*;
    - non-canonical representation of integers;
    - **not necessarily** a regular language.

s directed by β

- β belongs to $\mathbb{Q} \setminus \mathbb{N}$
  - linked to rational base number system;
  - non-canonical representation;
  - always a FLIP Language.
- β belongs to $\mathbb{N}$
  - linked to integer base *b*;
  - non-canonical representation of integers;
  - **not necessarily** a regular language.

- β is a Pisot number
  - linked to the NS built from the minimal polynomial of β;
  - non-canonical representation of integers;
  - **not necessarily** a regular language.
- β is neither rational nor a Pisot number
  - **not necessarily** linked to the NS built from the minimal polynomial of β.

Open problems

L: a regular language whose generating function is b^n
Is L directed by b?

L and K: two regular languages with the same generating function.
Are the paths associated with their signature bounded?

Which (regular) languages have sturmian words as their signature?
Is it linked to the NS whose base is the slope of this sturmian word?