A First Investigation of Sturmian Trees

Jean Berstel[2], Luc Boasson[1]

Olivier Carton[1], Isabelle Fagnot[2]

[1]LIAFA, CNRS
Université Paris 7

[2]IGM, CNRS
Université de Marne-la-Vallée

STACS'2007, Aachen

Outline

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$$x = a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ \ldots$$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$$x = \boxed{a}\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b \ldots$$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$$x = a\ \boxed{b}\ a\ a\ b\ a\ b\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ \ldots$$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$$x = \boxed{a\ b}\ a\ a\ b\ a\ b\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ \ldots$$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$$x = a\ \boxed{b\ a}\ a\ b\ a\ b\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b \ldots$$

Length: n	1	2	3	4	5
Factors	a	aa	aab	$aaba$	$aabaa$
	b	ab	aba	$abaa$	$aabab$
		ba	baa	$abab$	$abaab$
			bab	$baab$	$ababa$
				$baba$	$baaba$
					$babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$$x = a\ b\ \boxed{a\ a}\ b\ a\ b\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ \ldots$$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$x =$ a b a $\boxed{a\ b}$ a b a a b a a b a b a a b a b a a b …

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$x = \boxed{a\ b\ a}\ a\ b\ a\ b\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b \ldots$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$$x = a\ \boxed{b\ a\ a}\ b\ a\ b\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ \ldots$$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$$x = a\ b\ \boxed{a\ a\ b}\ a\ b\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ \ldots$$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$$x = a\ b\ a\ \boxed{a\ b\ a}\ b\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b \ldots$$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$$x = a\ b\ a\ a\ \boxed{b\ a\ b}\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ \ldots$$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$$x = \boxed{a\ b\ a\ a}\ b\ a\ b\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ \ldots$$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$$x = a\ \boxed{b\ a\ a\ b}\ a\ b\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b \ldots$$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$$x = a\ b\ \boxed{a\ a\ b\ a}\ b\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ \ldots$$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$$x = a\ b\ a\ \boxed{a\ b\ a\ b}\ a\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ \dots$$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Factors of an infinite word

A factor w of a word x is a finite word that occurs in x, that is, there are words u and y such that $x = uwy$.

Example (Fibonacci word)

$x = a\ b\ a\ a\ \boxed{b\ a\ b\ a}\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b\ a\ b\ a\ a\ b \ldots$

Length: n	1	2	3	4	5
Factors	a b	aa ab ba	aab aba baa bab	$aaba$ $abaa$ $abab$ $baab$ $baba$	$aabaa$ $aabab$ $abaab$ $ababa$ $baaba$ $babaa$
#	2	3	4	5	6

Sturmian words

Proposition (Hedlund & Morse)

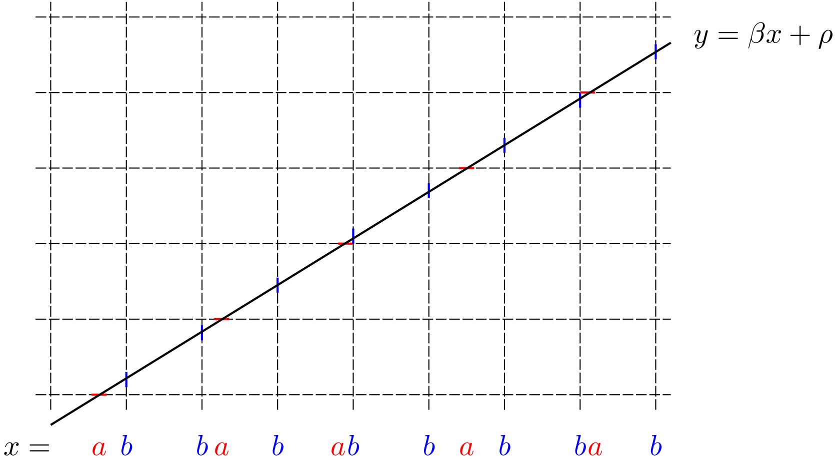
An infinite word x ultimately periodic iff there is an integer n such that x has at most n distinct factors of length n.

An infinite word x is Sturmian if the number of its factors of length n is $n+1$ for each n.
Sturmian words are non ultimately periodic words with the smallest complexity.

Example (Fibonacci word: $f_{n+2} = f_{n+1}f_n$)

$$f_0 = a$$
$$f_1 = ab$$
$$f_2 = aba$$
$$f_3 = abaab$$
$$f_\omega = abaababa\cdots$$

Characterization: cutting sequences

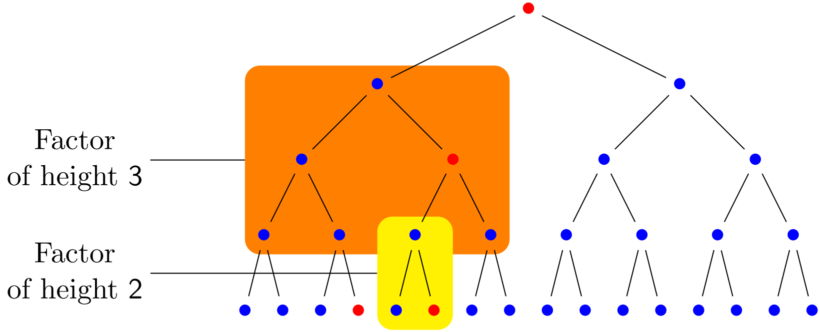


Theorem

A infinite word is Sturmian iff it is the cutting sequence of a straight line $y = \beta x + \rho$ with an irrational slope β.

Factor of a tree

A factor of height h of a tree t is a subtree of height h that occurs in t.



Sturmian tree

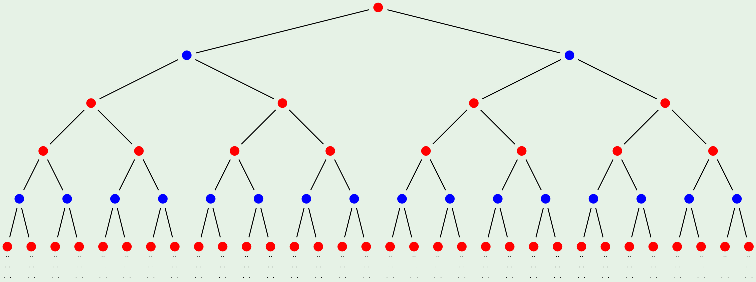
Proposition (Carpi et al)

A complete tree t is rational if there is some integer h such that t has at most h distinct factors of height h.

A tree is Sturmian if the number of its factors of height h is $h+1$ for each h.

Example (Easy one: uniform tree)

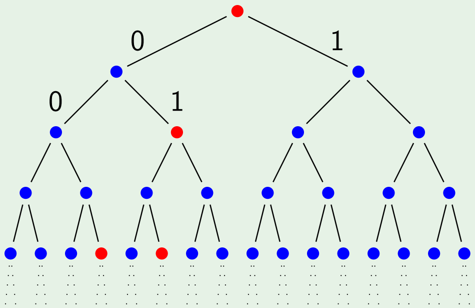
An Sturmian word $x = abaaba\ldots$ is repeated on each branch.



Example (Unexpected one: Dyck tree)

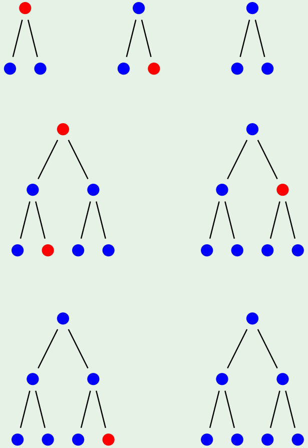
A node is • if it is a Dyck word over the alphabet $\{0, 1\}$.

The Dyck tree



$D_2^* = \{\varepsilon, 01, 0101, 0011, \ldots\}$

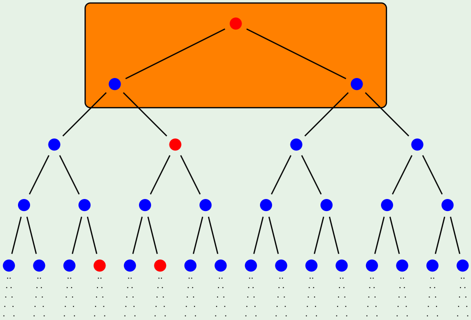
Its factors



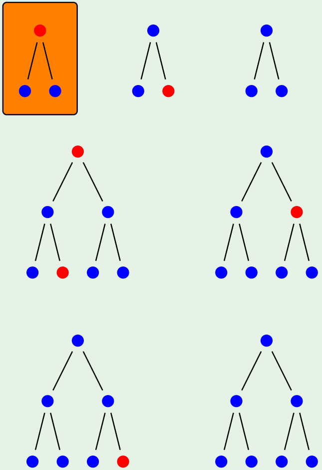
Example (Unexpected one: Dyck tree)

A node is • if it is a Dyck word over the alphabet $\{0, 1\}$.

The Dyck tree



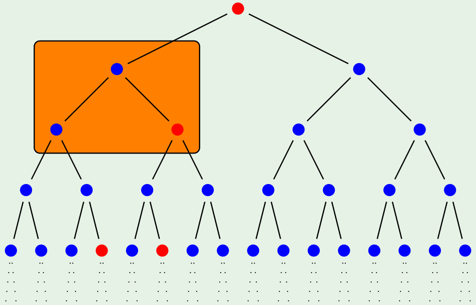
Its factors



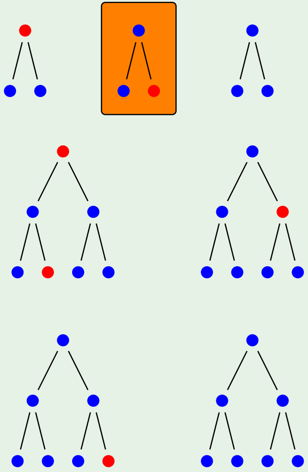
Example (Unexpected one: Dyck tree)

A node is • if it is a Dyck word over the alphabet $\{0, 1\}$.

The Dyck tree



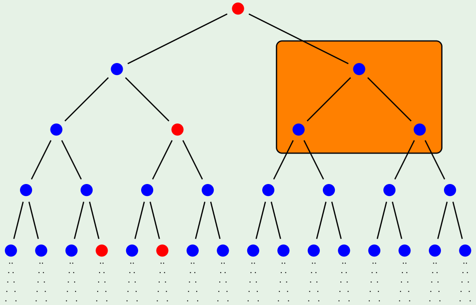
Its factors



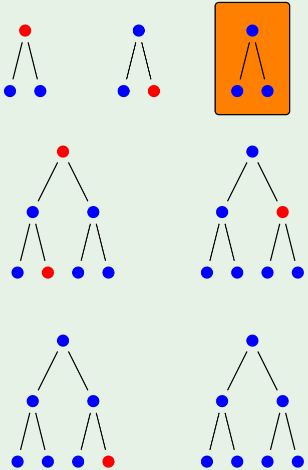
Example (Unexpected one: Dyck tree)

A node is ● if it is a Dyck word over the alphabet $\{0, 1\}$.

The Dyck tree



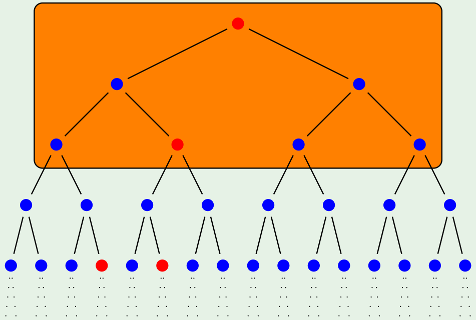
Its factors



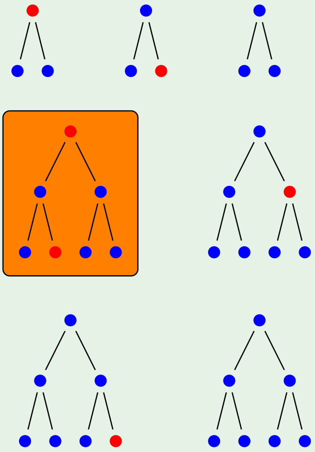
Example (Unexpected one: Dyck tree)

A node is $\bullet$ if it is a Dyck word over the alphabet $\{0, 1\}$.

The Dyck tree



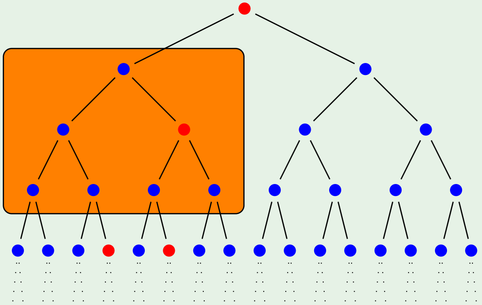
Its factors



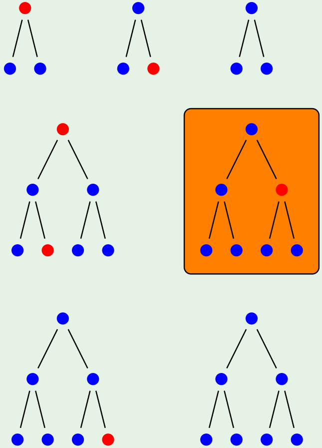
Example (Unexpected one: Dyck tree)

A node is $\bullet$ if it is a Dyck word over the alphabet $\{0, 1\}$.

The Dyck tree



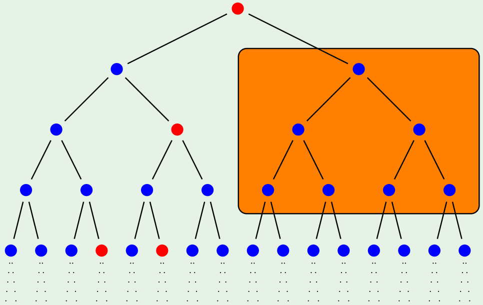
Its factors



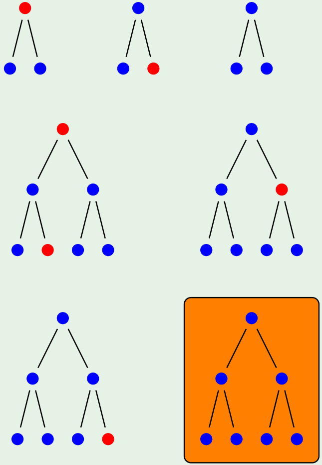
Example (Unexpected one: Dyck tree)

A node is ● if it is a Dyck word over the alphabet $\{0, 1\}$.

The Dyck tree



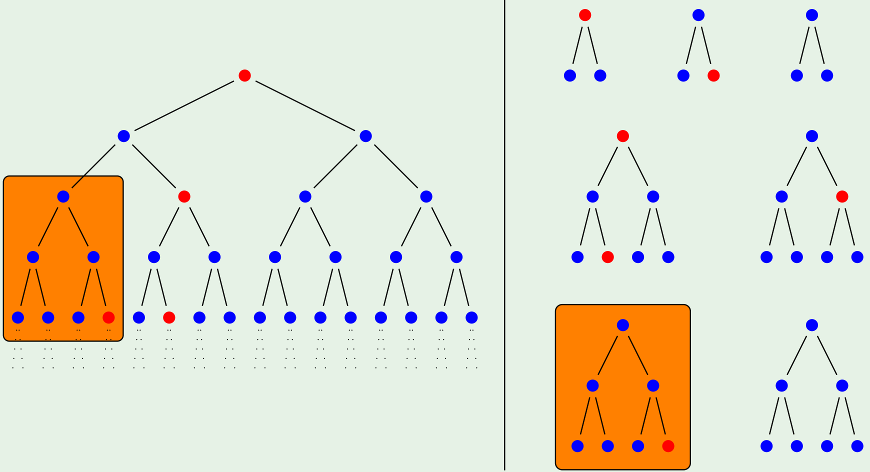
Its factors



Example (Unexpected one: Dyck tree)

A node is • if it is a Dyck word over the alphabet $\{0, 1\}$.

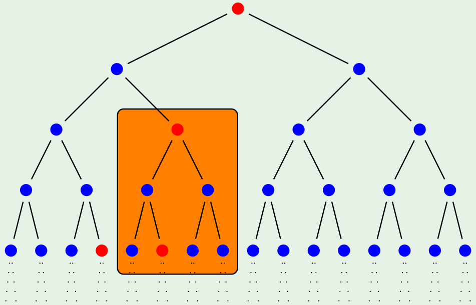
The Dyck tree | Its factors



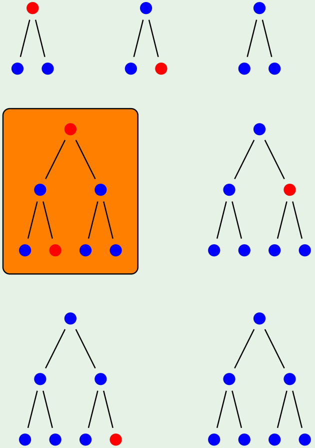
Example (Unexpected one: Dyck tree)

A node is • if it is a Dyck word over the alphabet $\{0, 1\}$.

The Dyck tree



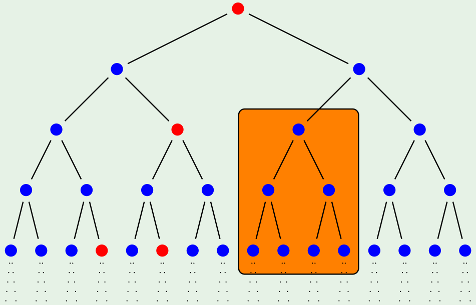
Its factors



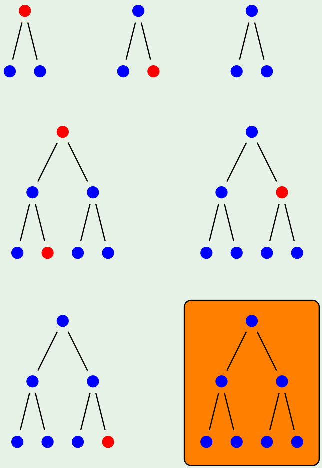
Example (Unexpected one: Dyck tree)

A node is $\bullet$ if it is a Dyck word over the alphabet $\{0, 1\}$.

The Dyck tree



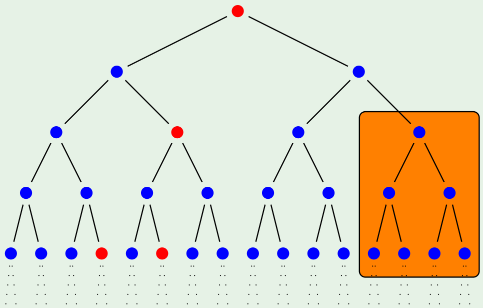
Its factors



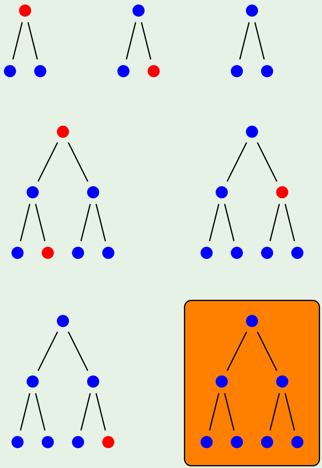
Example (Unexpected one: Dyck tree)

A node is $\bullet$ if it is a Dyck word over the alphabet $\{0, 1\}$.

The Dyck tree



Its factors



Slow automata

Let $\mathcal{A} = (Q, A, \cdot, q_0, F)$ be a (infinite) deterministic automaton. Define the Moore equivalence $\sim_n$ by induction.

$$q \sim_1 q' \iff (q \in F \Leftrightarrow q' \in F)$$
$$q \sim_{n+1} q' \iff (q \sim_n q') \text{ and } (\forall a \in A \ \ q \cdot a \sim_n q' \cdot a)$$

The relation $q \sim_n q'$ does not hold if there is a word w of length n such that $q \cdot w \in F$ and $q' \cdot w \notin F$ (or $q \cdot w \notin F$ and $q' \cdot w \in F$)

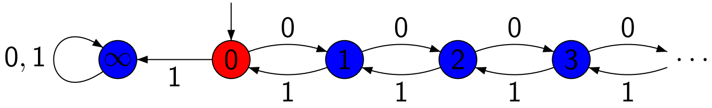
An infinite automaton is slow iff each equivalence $\sim_n$ has $n + 1$ classes.

Proposition

A tree t is Sturmian iff the minimal automaton of $t^{-1}(a)$ is slow.

Application to the Dyck tree

The minimal automaton of the Dyck language is the following.



The Moore equivalences of this automaton

$$\sim_1: \quad 0 \mid 1, 2, 3, 4, \ldots \infty$$
$$\sim_2: \quad 0 \mid 1 \mid 234 \ldots \infty$$
$$\sim_3: \quad 0 \mid 1 \mid 2 \mid 3, 4, \ldots \infty$$
$$\sim_4: \quad 0 \mid 1 \mid 2 \mid 3 \mid 4, \ldots \infty$$

Rank and degree

- A node is called irrational if the infinite subtree rooted in this node is not rational.
- The rank is the number of distinct rational subtrees.
- The degree is the number of branches of irrational nodes.

Examples

- The uniform tree has rank 0 and degree ∞.
- The Dyck tree has rank 1 and degree ∞.

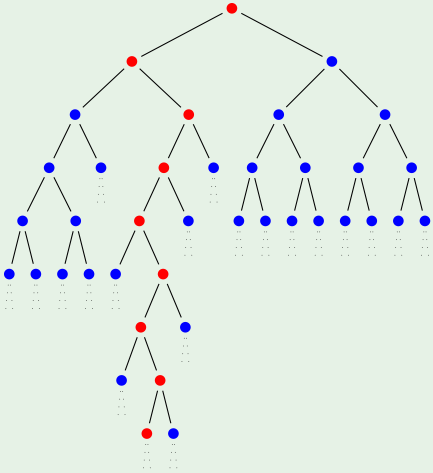
Results

degree	rank	
	finite	infinite
1	characterized	example later
≥ 2, finite	proved to be empty	example in full paper
infinite	example of Dyck tree	example in full paper

Characterization

Example (Indicator tree)

Take any Sturmian word (e.g. $01001010\cdots$) and distinguish the branch labeled by this word.



Rank 1

Degree 1

Yet another example

Example (Rank ∞ and degree 1)