Balanced Labeled Trees : Density, Complexity and Mechanicity

Nicolas Gast Bruno Gaujal

Laboratoire d'informatique de Grenoble

April 6, 2010

Sturmian Words: 3 equivalent definitions

Infinite word : 0010100100101001010100100...

- minimal complexity : $n+1$ factors of length n.
  example: 4 factors of length 3: 001, 010, 100 and 101.
- balanced : number of 1 only differ by 1 in factors of same length.
  - length 3: 1 or 2.
  - length 4: 1 or 2.
  - ...
- mechanical:
  - for all i: $w_i = \lfloor \alpha(i+1) + \theta \rfloor - \lfloor \alpha i + \theta \rfloor$
    or for all i: $w_i = \lceil \alpha(i+1) + \theta \rceil - \lceil \alpha i + \theta \rceil$

Problem

Can we extend theses notions to trees?

- complexity
- balanced
- mechanical

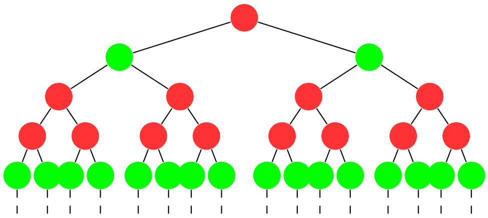
Previous Work

Definition (Berstel, Boasson, Carton and Fagnot, 2007)

A Sturmian tree is an infinite binary tree with $n+1$ subtrees of height n.

(subtree = complete subtree)

Simple example:

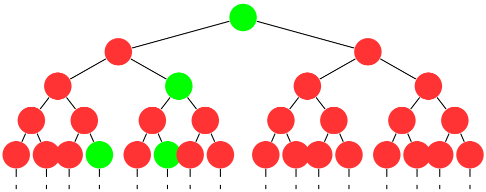


Example: The uniform tree corresponding to 0100101 . . .

Properties

Let $L \subset \{0;1\}^*$. We define a tree $T(L)$:
A node w of $T(L)$ has a label 1 *iff* $w \in L$.

- The sturmian property of L can be viewed as a property of the language L.
- Lots of interesting examples.



Example: Dyck Tree $\{\epsilon, ab, aabb, abab, \dots\}$

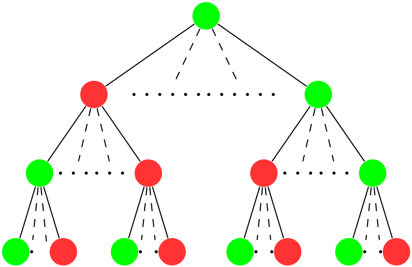
But

- the *balanced* property is lost (important in optimization problems)
- no simple equivalent characterization

Our trees are:

- rooted
- labeled by 0 or 1
- infinite
- constant degree $d \geq 2$ ($d = 1$ is words)
- Non-ordered ($\neq$ Original definition of Sturmian Trees)

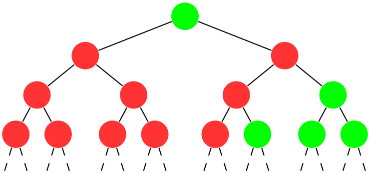
Each node has exactly d children.



The Set of Trees

Our trees are:

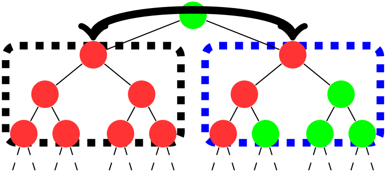
- rooted
- labeled by 0 or 1
- infinite



- constant degree $d \geq 2$
  ($d = 1$ is words)
- Non-ordered
  ($\neq$ Original definition of Sturmian Trees)

The Set of Trees

Our trees are:

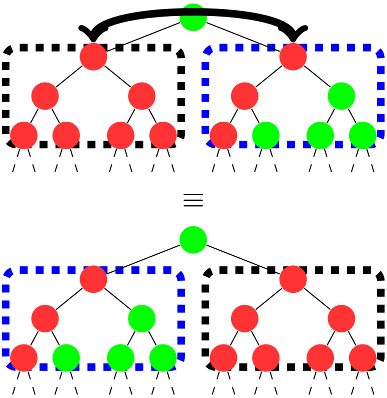


- rooted
- labeled by 0 or 1
- infinite
- constant degree $d \geq 2$
  ($d = 1$ is words)
- Non-ordered
  ($\neq$ Original definition of Sturmian Trees)

The Set of Trees

Our trees are:

- rooted
- labeled by 0 or 1
- infinite
- constant degree $d \geq 2$ ($d = 1$ is words)
- Non-ordered (≠ Original definition of Sturmian Trees)

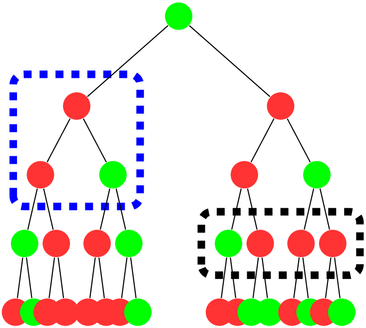


What are Subtrees?

We define:

- Subtree of height n
  $\frac{d^n-1}{d-1}$ nodes
- Truncated subtree of width k and height n
  $\frac{d^n-d^k}{d-1}$ nodes

- Blue: Subtree of height 2
- Black: Truncated subtree of height 1 and width 2



Density - 1/2

Density of a word: simple definition.

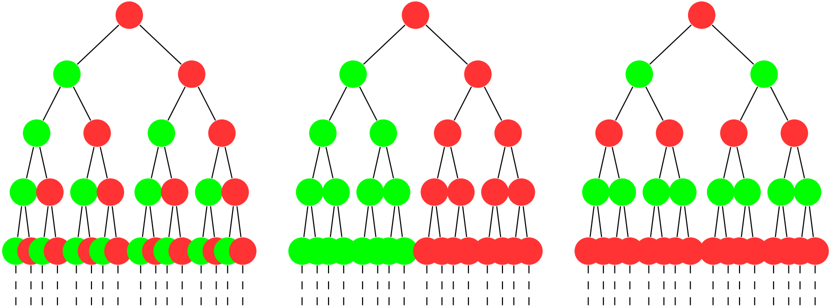
- $w = w_1 w_2 w_3 \ldots \ (w_i \in \{0;1\})$
- $\text{density} = \lim \frac{\sum_{i \leq n} w_i}{n}$

Density - 1/2

Density of a word: simple definition.

- $w = w_1 w_2 w_3 \dots \ (w_i \in \{0;1\})$
- $\mathrm{density} = \lim \frac{\sum_{i \leq n} w_i}{n}$

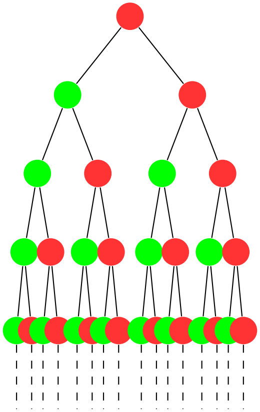
Do theses trees have a density?



Density - 2/2

Tree: two definitions (to be improved)

- Density of a finite subtree = average number of 1.
- If d_n is the density of the rooted subtree of height n:
  - density $= \lim_n d_n$
  - average density $= \lim_n \frac{1}{n} \sum_{k=1}^{n} d_k$

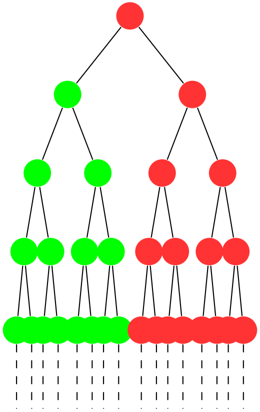


Density: $1/2$.

Density - 2/2

Tree: two definitions (to be improved)

- Density of a finite subtree = average number of 1.
- If d_n is the density of the rooted subtree of height n:
  - density $= \lim_n d_n$
  - average density $= \lim_n \frac{1}{n} \sum_{k=1}^{n} d_k$

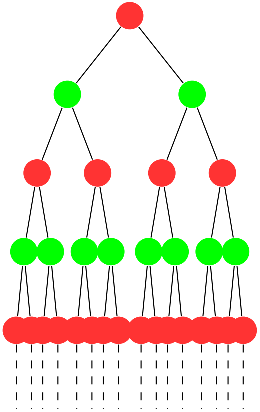


Density: 1/2
(for this rooted definition)

Density - 2/2

Tree: two definitions (to be improved)

- Density of a finite subtree = average number of 1.
- If d_n is the density of the rooted subtree of height n:
  - density $= \lim_n d_n$
  - average density $= \lim_n \frac{1}{n} \sum_{k=1}^{n} d_k$



No density.
Average density: $1/2$.

Rational Words

Let w be an infinite word on $\{0, 1\}$ and $F_w(n)$ be the number of factor of w of size n. A word is rational if $F_w(n)$ is bounded (or equivalently if $F_w(n) \leq n$ for any n).

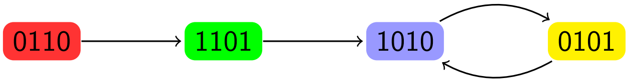
Rational Words

Let w be an infinite word on $\{0, 1\}$ and $F_w(n)$ be the number of factor of w of size n. A word is rational if $F_w(n)$ is bounded (or equivalently if $F_w(n) \leq n$ for any n).

0110 1101 1010 0101

Rational Words

Let w be an infinite word on $\{0,1\}$ and $F_w(n)$ be the number of factor of w of size n. A word is rational if $F_w(n)$ is bounded (or equivalently if $F_w(n) \leq n$ for any n).



Rational Words

Let w be an infinite word on $\{0, 1\}$ and $F_w(n)$ be the number of factor of w of size n. A word is rational if $F_w(n)$ is bounded (or equivalently if $F_w(n) \leq n$ for any n).

$$\begin{tikzcd} 0110 \arrow[r] & 1101 \arrow[r] & 1010 \arrow[r, bend left] & 0101 \arrow[l, bend left] \end{tikzcd}$$

Lots of properties:

- w is ultimately periodic: $w = xy^{\infty}$: $\underbrace{011}_{x}\,\underbrace{0101010101\ldots}_{y^{\infty}}$
- density (here: $1/2$)

Rational Words

Let w be an infinite word on $\{0,1\}$ and $F_w(n)$ be the number of factor of w of size n. A word is rational if $F_w(n)$ is bounded (or equivalently if $F_w(n) \leq n$ for any n).

$$\begin{tikzcd} 0110 \arrow[r] & 1101 \arrow[r] & 1010 \arrow[r, bend left] & 0101 \arrow[l, bend left] \end{tikzcd}$$

Lots of properties:

- w is ultimately periodic: $w = xy^\infty$: $\underbrace{011}_{x}\underbrace{0101010101\ldots}_{y^\infty}$
- density (here: $1/2$)

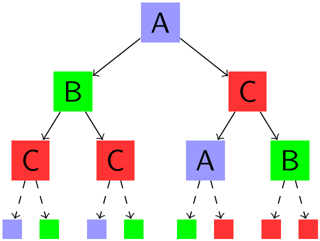
Recall: a sturmian word has exactly $n+1$ words of size n.
Thus sturmian words are aperiodic words of minimal complexity.

Rational Trees: Definition

We call $P(n)$ = number of subtrees of height n.

Rational Trees: 3 equivalent definitions:

- $P(n)$ bounded.
- $\exists n / P(n) = P(n+1)$
- $\exists n / P(n) \leq n$.

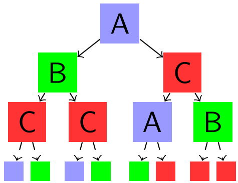


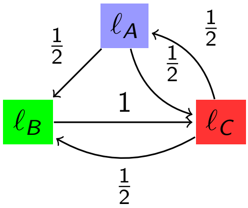
Theorem

- *A rational Tree has an average density α which is rational.*

  α *is not necessary a density but:*

- *If the associated Markov chain is irreducible and aperiodic then there exists a density.*

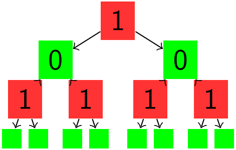


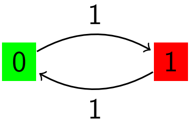


Example: A rational tree and its associated Markov chain

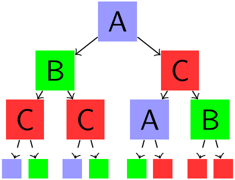
Example of Density

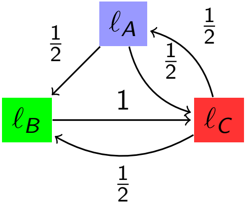
- Periodic $=$ average density $d_{\text{average}} = \frac{1}{2}$





- Aperiodic : density $d = \frac{2}{9}\ell_A + \frac{1}{3}\ell_B + \frac{4}{9}\ell_C$



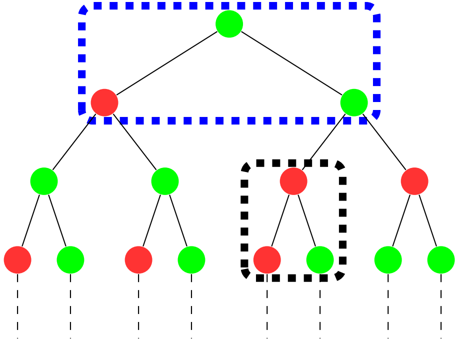


Balanced Trees and Strongly Balanced Trees

- Balanced tree: number of 1 in subtrees of height n only differ by 1.
- Strongly balanced tree: same property with truncated subtrees of height n and width k.

Balanced Trees and Strongly Balanced Trees

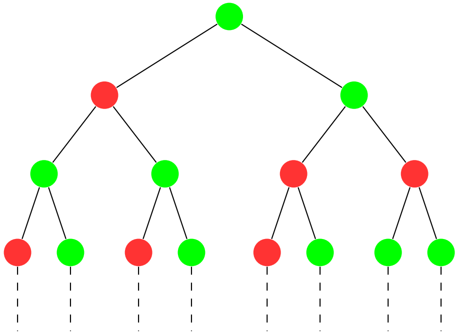
- Balanced tree: number of 1 in subtrees of height n only differ by 1.
- Strongly balanced tree: same property with truncated subtrees of height n and width k.



Example: Balanced tree not strongly balanced

Balanced Trees and Strongly Balanced Trees

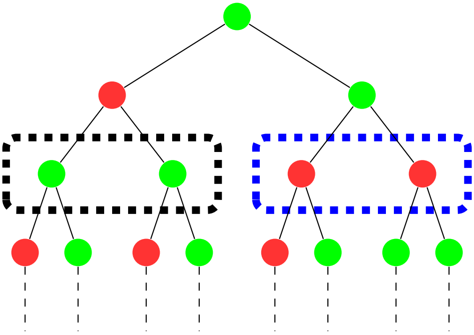
- Balanced tree: number of 1 in subtrees of height n only differ by 1.
- Strongly balanced tree: same property with truncated subtrees of height n and width k.



Example: Balanced tree not strongly balanced

Balanced Trees and Strongly Balanced Trees

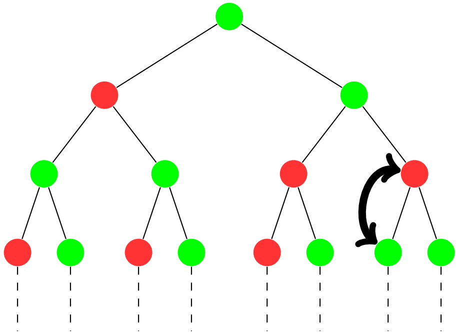
- Balanced tree: number of 1 in subtrees of height n only differ by 1.
- Strongly balanced tree: same property with truncated subtrees of height n and width k.



Example: Balanced tree not strongly balanced

Balanced Trees and Strongly Balanced Trees

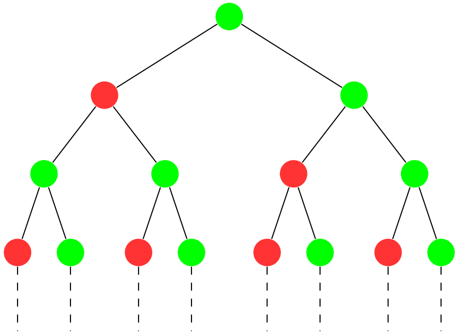
- Balanced tree: number of 1 in subtrees of height n only differ by 1.
- Strongly balanced tree: same property with truncated subtrees of height n and width k.



Example: Balanced tree not strongly balanced

Balanced Trees and Strongly Balanced Trees

- Balanced tree: number of 1 in subtrees of height n only differ by 1.
- Strongly balanced tree: same property with truncated subtrees of height n and width k.



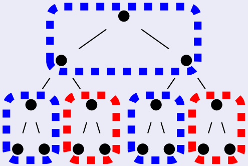
Example: Strongly balanced tree

Density of a Balanced Tree

Theorem

- *A balanced tree has a density.*

Sketch of the proof.

1. A tree of height n has a density α_n or $\alpha_n + \frac{d-1}{d^n-1}$
2. 

   If blue has a density α_2 and red $\alpha_2 + \frac{1}{3}$ then $\alpha_2 \leq \alpha_4 \leq \alpha_2 + \frac{1}{3}$
   $\alpha_n \leq \alpha_{2n} + \alpha_n + \epsilon$
3. Take limit

$\square$

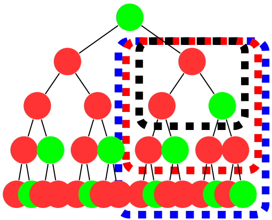
Mechanical Trees

We want to write a numerical formula to build trees of density α.

- A subtree of height n has $\frac{d^n-1}{d-1}$ nodes.
- We want density α

We want to write a numerical formula to build trees of density α.

- A subtree of height n has $\frac{d^n-1}{d-1}$ nodes.
- We want density α



$\alpha = 0.3$, $\phi = 0.55$.

n	$(2^n - 1)\alpha + \phi$
1	0.85
2	1.45
3	2.65
4	5.05

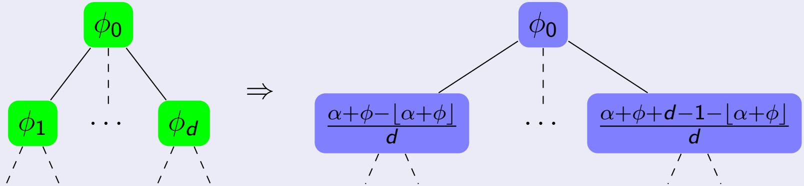
Definition

A tree is mechanical of density α if for all node i, there exists a phase $\phi_i \in [0;1)$ such that the number of 1 in a subtree of height n and root i is $\lfloor \frac{d^n-1}{d-1}\alpha + \phi_i \rfloor$ (resp. for all i: $\lceil \frac{d^n-1}{d-1}\alpha + \phi_i \rceil$)

Uniqueness of a mechanical Tree

Theorem

- *There exists a unique mechanical tree if* (α, ϕ_0) *is fixed.*



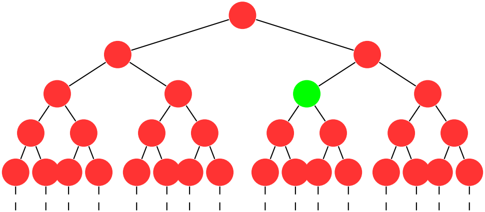
- *The phase of the root* ϕ_0 *is unique for almost all* α.

- Relations between Mechanical and Balanced trees?
- Limit of the equivalences

Equivalences between Definitions

Theorem (Mechanical $\sim$ strongly balanced)

- *A mechanical tree is strongly balanced*
- *A strongly balanced tree with density different from* $\frac{p(d-1)}{d^n-d^w}$ *is mechanical*
- *A strongly balanced tree with density* $\frac{p(d-1)}{d^n-d^w}$ *is ultimately mechanical.*



Example: Ultimately mechanical tree

Sketch of Proof

Mechanical implies strongly balanced.

The number of 1 in a truncated subtree of height n and width k is bounded by $\lfloor \frac{d^n-d^k}{d-1}\alpha \rfloor$ and $\lfloor \frac{d^n-d^k}{d-1}\alpha \rfloor + 1$ □

Strongly Balanced implies mechanical.

$\forall \tau \in [0;1)$, if h_n is the number of 1 in the subtree of size n, at least one of these properties is true:

1. for all n: $h_n \leq \lfloor \frac{d^n-1}{d-1}\alpha + \tau \rfloor$,
2. for all n: $h_n \geq \lfloor \frac{d^n-1}{d-1}\alpha + \tau \rfloor$.

We choose ϕ the maximal τ such that 1 is true. □

Mechanical Trees are Sturmian

Theorem

- *A mechanical tree is rational or sturmian.*
- *A mechanical tree of density different from* $\frac{p(d-1)}{d^n - d^w}$ *is a sturmian tree: it has* $n+1$ *subtrees of height* n.

Mechanical Trees are Sturmian

Theorem

- *A mechanical tree is rational or sturmian.*
- *A mechanical tree of density different from* $\frac{p(d-1)}{d^n-d^w}$ *is a sturmian tree: it has* $n+1$ *subtrees of height* n*.*

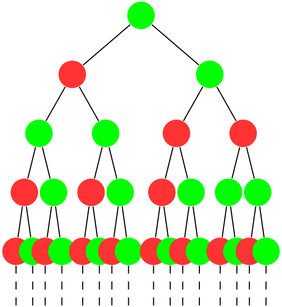
Proof.

- A subtree of height n depends only on its phase
- In fact, it depends on $(\alpha+\phi,\ldots,\frac{d^n-1}{d-1}\alpha+\phi)$ which takes $n+1$ values when $\phi\in[0;1)$.

$\square$

Limit of the Equivalences

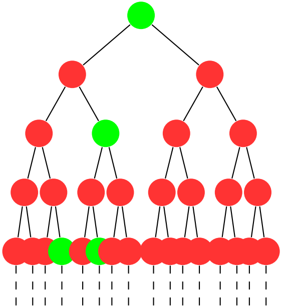
- Balanced $\not\Rightarrow$ strongly balanced (no matter whether the density is rational or not).
- Sturmian $\not\Rightarrow$ balanced.
- Irrational Balanced tree $\not\Rightarrow$ sturmian.



Example: Balanced tree not str. bal.

Limit of the Equivalences

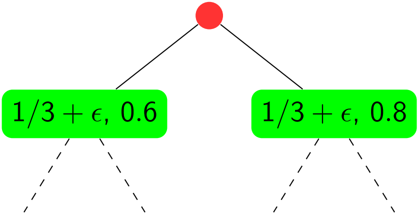
- Balanced $\not\Rightarrow$ strongly balanced (no matter whether the density is rational or not).
- Sturmian $\not\Rightarrow$ balanced.
- Irrational Balanced tree $\not\Rightarrow$ sturmian.



Example: Dyck Tree

Limit of the Equivalences

- Balanced $\not\Rightarrow$ strongly balanced (no matter whether the density is rational or not).
- Sturmian $\not\Rightarrow$ balanced.
- Irrational Balanced tree $\not\Rightarrow$ sturmian.



Example: Balanced tree non sturmian

- Links between strongly balanced trees and rational trees
  - How to test if a rational tree is strongly balanced?
- Compute the number of mechanical trees of size n.

Links between rational and Strongly Balanced Tree

1. A strongly balanced tree with density different from $\frac{p(d-1)}{d^n-d^w}$ is sturmian.
2. A strongly balanced tree with density $\frac{p(d-1)}{d^n-d^w}$ is rational.

Thus:

Corollary

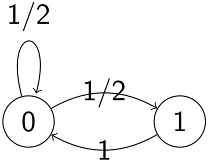
There is no mechanical rational tree of density different from $\frac{p(d-1)}{d^n-d^w}$.

Although that for all p, q:

- there exists rational trees of density p/q
- there exists mechanical trees of density p/q

Testing if a Tree is Strongly Balanced - 1/2

Problem: given an *irreducible* Markov chain (=finite representation of a tree) corresponding to a rational tree, can we say if this tree balanced?



Example: Is the corresponding tree strongly balanced?

Remark: when the Chain is irreducible, a rational tree is balanced *iif* it is mechanical.

Testing if a Rational Tree is Strongly Balanced - 2/2

Algorithm:

1. Compute the density α.
2. If α is not some $\frac{p(d-1)}{d^n - d^w}$ then the tree is not balanced
3. Otherwise, we just have to test the balance property on the truncated subtrees of height at most n.

Testing if a Rational Tree is Strongly Balanced - 2/2

Algorithm:

1. Compute the density α.
2. If α is not some $\frac{p(d-1)}{d^n - d^w}$ then the tree is not balanced
3. Otherwise, we just have to test the balance property on the truncated subtrees of height at most n.

Complexity.

If the Markov Chain has n states, the complexity is:

1. compute the stationary distribution: $O(n^3)$
2. We can prove that if the tree is balanced, then $\alpha = \frac{p(d-1)}{d^n - d^w}$ where n is the number of the states of the Markov Chain: $O(n)$ values for w to test.
3. For all height $n', w' < n$, there are at most n subtrees, so $O(n^3)$ trees to test (with 2^n nodes at most).

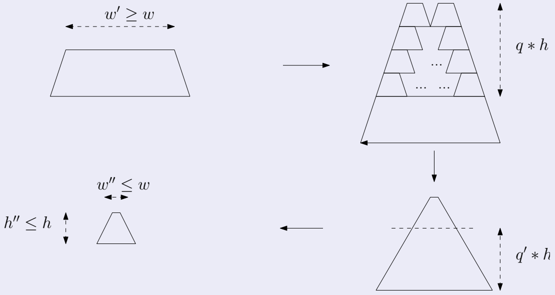
$\square$

Algorithm:

1. Compute the density α.
2. If α is not some $\frac{p(d-1)}{d^n - d^w}$ then the tree is not balanced
3. Otherwise, we just have to test the balance property on the truncated subtrees of height at most n.

Proof of line 3.

Each truncated subtree of size (h, w) has exactly p 1.



□

Number of Mechanical Trees of size n

Recall: number of sturmian words of size n is $1 + \sum_1^n (n-i+1)\phi(i) \sim \frac{n^3}{\pi}$ where ϕ is the Euler function (number of natural integers less than n and co-prime with n).

Number of Mechanical Trees of size n

Recall: number of sturmian words of size n is $1 + \sum_1^n (n - i + 1)\phi(i) \sim \frac{n^3}{\pi}$ where ϕ is the Euler function (number of natural integers less than n and co-prime with n).

Theorem (number c_n of mechanical subtrees of size n)

For all n: $2^n \leq c_n \leq n2^n$

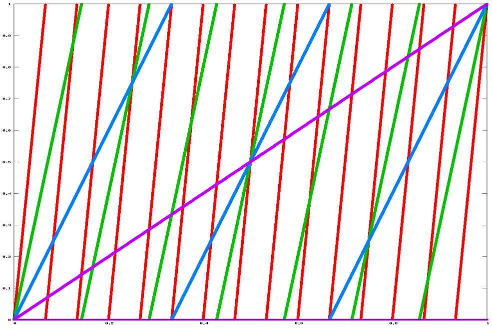
Number of Mechanical Trees of size n

Recall: number of sturmian words of size n is $1 + \sum_1^n (n - i + 1)\phi(i) \sim \frac{n^3}{\pi}$ where ϕ is the Euler function (number of natural integers less than n and co-prime with n).

Theorem (number c_n of mechanical subtrees of size n)

For all n: $2^n \leq c_n \leq n2^n$

$c_n =$ number of tuples $((2^1 - 1)\alpha + \phi, \ldots, (d^n - 1)\alpha + \phi)$



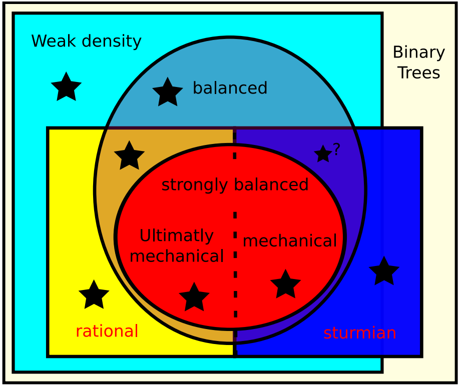
- Draw $\{(\alpha, \phi)$ such that $(2^n - 1)\alpha + \phi \in \mathbb{N}\}$
- Compute the number of areas.
- Same kind of estimations for degree d.

Conclusion

- Unordered definition better here
- Constructive definition
- Strong inclusions
- Good characterization

but:

- better definition for density
- balanced definition?
- study optimization issues



★ = we know a counter-example

★? = we think there is a counter-example

Thanks for your attention!