Growing binary trees

Khaydar Nurligareev (with Antoine Genitrini)

LIP6, Sorbonne University

Journées Graphes et Algorithmes 2025

IPGP, Paris

November 18, 2025

Settings

Vertex types

- ● internal node
- ○ anchor (active leaf)
- □ leaf (dead leaf)

Settings

Vertex types

- $\bullet$ internal node
- $\circ$ anchor (active leaf)
- $\square$ leaf (dead leaf)

$t = 0$ $\circ$

Growing process

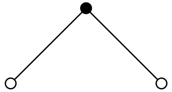
- At the beginning ($t = 0$), our tree is an anchor $\circ$

Settings

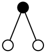
Vertex types

- • internal node
- ○ anchor (active leaf)
- □ leaf (dead leaf)

$t = 1$



Growing process

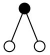
- At the beginning ($t = 0$), our tree is an anchor ○
- At any moment $t \geqslant 1$, replace each anchor ○ by
  - a leaf □
  - or a subtree 

Settings

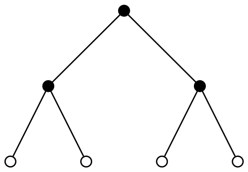
Vertex types

- ● internal node
- ○ anchor (active leaf)
- □ leaf (dead leaf)

Growing process

- At the beginning ($t = 0$), our tree is an anchor ○
- At any moment $t \geqslant 1$, replace each anchor ○ by
  - a leaf □
  - or a subtree 

$t = 2$

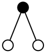


Settings

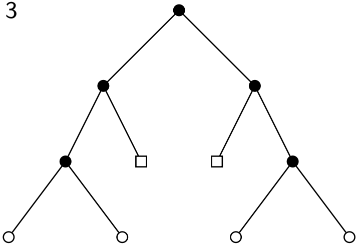
Vertex types

- ● internal node
- ○ anchor (active leaf)
- □ leaf (dead leaf)

Growing process

- At the beginning ($t = 0$), our tree is an anchor ○
- At any moment $t \geqslant 1$, replace each anchor ○ by
  - a leaf □
  - or a subtree 

$t = 3$

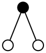


Settings

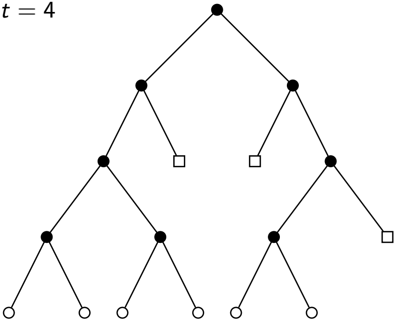
Vertex types

- $\bullet$ internal node
- $\circ$ anchor (active leaf)
- $\square$ leaf (dead leaf)

Growing process

- At the beginning ($t = 0$), our tree is an anchor $\circ$
- At any moment $t \geqslant 1$, replace each anchor $\circ$ by
  - a leaf $\square$
  - or a subtree 

$t = 4$

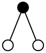


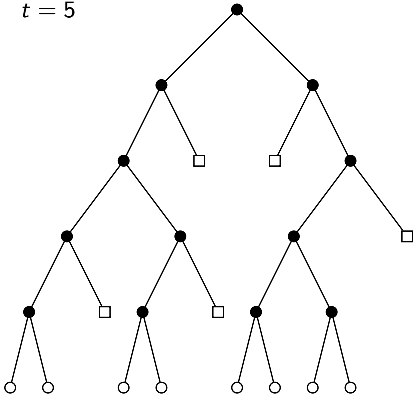
Settings

Vertex types

- ● internal node
- ○ anchor (active leaf)
- □ leaf (dead leaf)

Growing process

- At the beginning ($t = 0$), our tree is an anchor ○
- At any moment $t \geqslant 1$, replace each anchor ○ by
  - a leaf □
  - or a subtree 

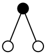


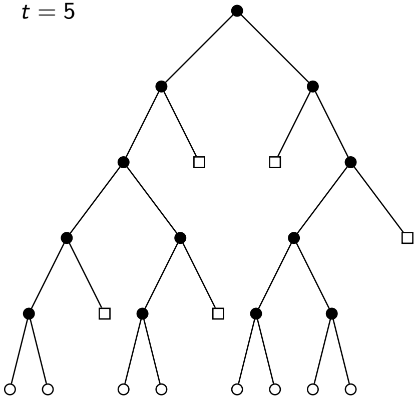
Settings

Vertex types

- ● internal node
- ○ anchor (active leaf)
- □ leaf (dead leaf)

Growing process

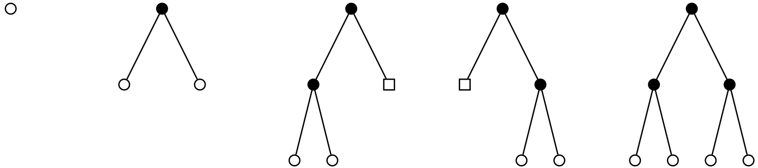
- At the beginning ($t = 0$), our tree is an anchor ○
- At any moment $t \geqslant 1$, replace each anchor ○ by
  - a leaf □
  - or a subtree 



Studied objects: active trees (i.e. trees that have anchors)

Counting

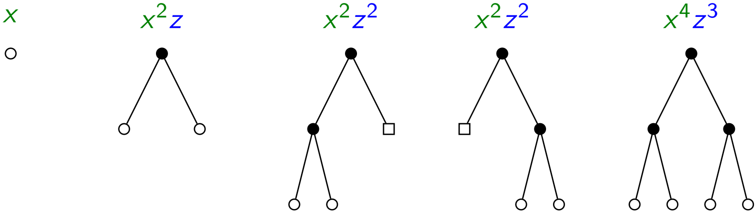
$$t_{n,m} = \{\text{active trees with } n \text{ internal nodes } m \text{ anchors}\}$$



$t_{0,1} = 1$ $t_{1,2} = 1$ $t_{2,2} = 2$ $t_{3,4} = 1$

Counting

$$t_{n,m} = \{\text{active trees with } n \text{ internal nodes } m \text{ anchors}\}$$



$t_{0,1} = 1$ $\quad$ $t_{1,2} = 1$ $\quad$ $t_{2,2} = 2$ $\quad$ $t_{3,4} = 1$

Marking variables:

- x marks anchors
- z marks internal nodes

$$T(x,z) = \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} t_{n,m} x^m z^n$$

$$T(x,z) = x + x^2 z + 2x^2 z^2 + \ldots$$

Relations

- Growing process replacement:

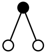
$$x \mapsto 1 \qquad\qquad x \mapsto zx^2$$

- Equation:

$$T(x,z) = x + T(1 + zx^2, z) - T(1, z)$$

Relations

- Growing process replacement:

$$\circ \mapsto \square \qquad\qquad \circ \mapsto$$


$$x \mapsto 1 \qquad\qquad x \mapsto zx^2$$

- Equation:

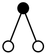
$$T(x,z) = x + T(1+zx^2, z) - C(z)$$

  - C_n are Catalan numbers
  - $C(z) = 1 + zC^2(z)$

$$\sum_{m=1}^{\infty} t_{n,m} = C_n$$

Relations

- Growing process replacement:

$$\circ \mapsto \square \qquad\qquad \circ \mapsto$$ 

$$x \mapsto 1 \qquad\qquad x \mapsto zx^2$$

- Equation:

$$T(x,z) = x + T(1+zx^2, z) - C(z)$$

  - C_n are Catalan numbers
  - $C(z) = 1 + zC^2(z)$

$$\sum_{m=1}^{\infty} t_{n,m} = C_n$$

- Recurrent relation $(n, k > 0)$:

$$t_{n,2k-1} = 0 \qquad \text{and} \qquad t_{n,2k} = \sum_{m=k}^{\infty} \binom{m}{k} t_{n-k,m}$$

Mandelbrot Polynomials

- Define

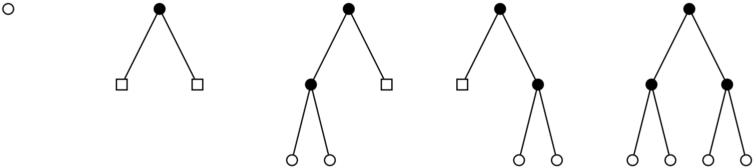
$$p_0(x,z) = x, \qquad p_{n+1}(x,z) = 1 + zp_n^2(x,z)$$

Mandelbrot Polynomials

- Define

$$p_0(x, z) = x, \qquad p_{n+1}(x, z) = 1 + z p_n^2(x, z)$$

- $p_n(x, z)$ counts binary trees:
  - of height at most n,
  - anchors are at level n,
  - leaves are at levels $k < n$



$$p_2(x, z) = 1 + z + 2x^2z^2 + x^4z^3$$

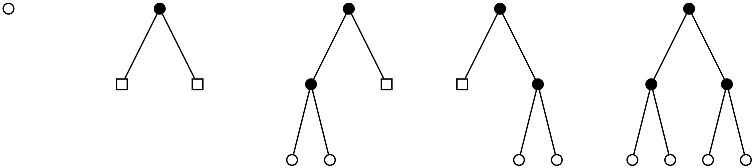
Mandelbrot Polynomials

- Define

$$p_0(x,z) = x, \qquad p_{n+1}(x,z) = 1 + zp_n^2(x,z)$$

- $p_n(x,z)$ counts binary trees:
  - of height at most n,
  - anchors are at level n,
  - leaves are at levels $k < n$

for $k > n$:
$[z^k]p_n(x,z) = C_k$



$$p_2(x,z) = 1 + z + 2x^2z^2 + x^4z^3$$

Mandelbrot Polynomials

- Define

$$p_0(x,z) = x, \qquad p_{n+1}(x,z) = 1 + zp_n^2(x,z)$$

- $p_n(x,z)$ counts binary trees:
  - of height at most n,
  - anchors are at level n,
  - leaves are at levels $k < n$

for $k > n$:
$[z^k]p_n(x,z) = C_k$

- Define

$$q_n(x,z) = zp_n(x,z)$$

- $q_n(1,z)$ are known as Mandelbrot Polynomials

Mandelbrot Polynomials

- Define

$$p_0(x,z) = x, \qquad p_{n+1}(x,z) = 1 + zp_n^2(x,z)$$

- $p_n(x,z)$ counts binary trees:
  - of height at most n,
  - anchors are at level n,
  - leaves are at levels $k < n$

for $k > n$:
$[z^k]p_n(x,z) = C_k$

- Define

$$q_n(x,z) = zp_n(x,z)$$

- $q_n(1,z)$ are known as Mandelbrot Polynomials

- Corollary: for $k < n$, $[z^{k+1}]q_n(1,z) = C_k$

Cumulative value of anchors

- Denote

$$\widetilde{T}(x,z) = \frac{\partial T}{\partial x}(x,z)$$

- Equation:

$$\widetilde{T}(x,z) = 1 + 2xz\,\widetilde{T}(1+zx^2, z)$$

Cumulative value of anchors

- Denote

$$\widetilde{T}(x,z) = \frac{\partial T}{\partial x}(x,z)$$

- Equation:

$$\widetilde{T}(x,z) = 1 + 2xz\,\widetilde{T}(1+zx^2, z)$$

- Relation:

$$\widetilde{T}(x,z) = 1 + \sum_{k=1}^{\infty} (2z)^k \prod_{\ell=0}^{k-1} p_\ell(x,z)$$

Cumulative value of anchors

- Denote

$$\widetilde{T}(x,z) = \frac{\partial T}{\partial x}(x,z)$$

- Equation:

$$\widetilde{T}(x,z) = 1 + 2xz\,\widetilde{T}(1+zx^2, z)$$

- Relation:

$$\widetilde{T}(x,z) = 1 + \sum_{k=1}^{\infty}(2z)^k \prod_{\ell=0}^{k-1} p_\ell(x,z)$$

- In terms of Mandelbrot polynomials:

$$\sum_{n=0}^{\infty}\sum_{m=1}^{\infty} m t_{m,n} z^n = 1 + \sum_{k=1}^{\infty} 2^k \prod_{\ell=0}^{k-1} q_\ell(1,z)$$

Anchor distributions

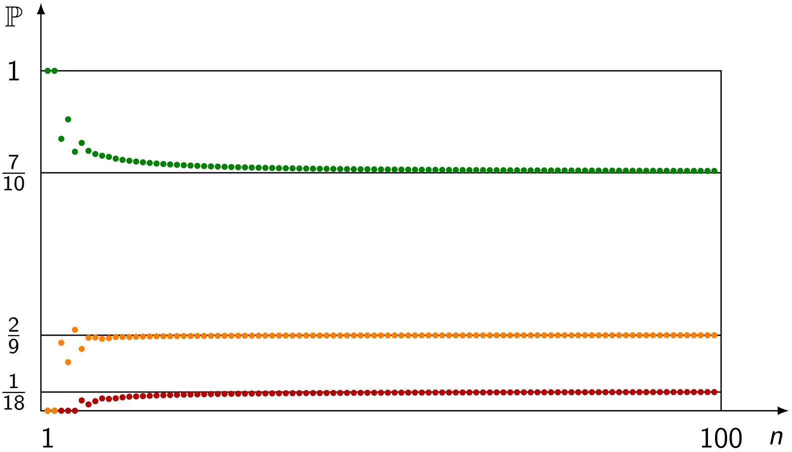
n	1	2	3	4	5	6	7	8	9	10	11
$t_{n,2}$	1	2	4	12	32	104	328	1 080	3 648	12 544	43 600
$t_{n,4}$	0	0	1	2	10	24	92	308	1 028	3 584	12 736
$t_{n,6}$	0	0	0	0	0	4	8	40	176	584	2 144
$t_{n,8}$	0	0	0	0	0	0	1	2	10	84	282
$t_{n,10}$	0	0	0	0	0	0	0	0	0	0	24
$t_{n,12}$	0	0	0	0	0	0	0	0	0	0	0
C_n	1	2	5	14	42	132	429	1 430	4 862	16 796	58 786

It looks like eventually

- $\frac{t_{n,2}}{C_n}$ is decreasing,
- $\frac{t_{n,2k}}{C_n}$ is increasing for $k > 1$,

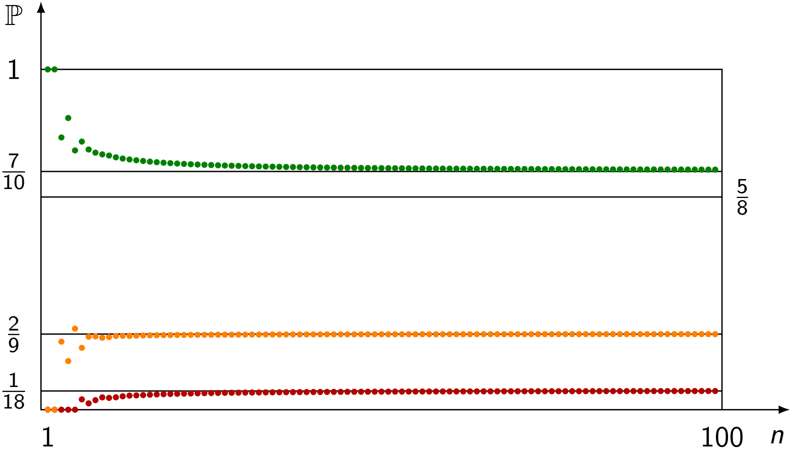
and there are some limits.

Proportions of the first three lines



Note that $\frac{t_{200,4}}{C_{200}} > \frac{2}{9}$ and $\frac{t_{200,6}}{C_{200}} > \frac{1}{18}$

Proportions of the first three lines



Proposition:

$$\liminf_{n\to\infty} \frac{t_{n,2}}{C_n} \geqslant \frac{5}{8}$$

Column nonzero values

n	1	2	3	4	5	6	7	8	9	10	11
$t_{n,2}$	1	2	4	12	32	104	328	1 080	3 648	12 544	43 600
$t_{n,4}$	0	0	1	2	10	24	92	308	1 028	3 584	12 736
$t_{n,6}$	0	0	0	0	0	4	8	40	176	584	2 144
$t_{n,8}$	0	0	0	0	0	0	1	2	10	84	282
$t_{n,10}$	0	0	0	0	0	0	0	0	0	0	24

Column nonzero values

n	1	2	3	4	5	6	7	8	9	10	11
$t_{n,2}$	1	2	4	12	32	104	328	1 080	3 648	12 544	43 600
$t_{n,4}$	0	0	1	2	10	24	92	308	1 028	3 584	12 736
$t_{n,6}$	0	0	0	0	0	4	8	40	176	584	2 144
$t_{n,8}$	0	0	0	0	0	0	1	2	10	84	282
$t_{n,10}$	0	0	0	0	0	0	0	0	0	0	24

- Define $$a_n = \max\{k \colon t_{n,2k} > 0\}$$

n	1	2	3	4	5	6	7	8	9	10
a_n	1	1	2	2	2	3	4	4	4	4
a_{n+10}	5	6	6	7	8	8	8	8	8	9
a_{n+20}	10	10	11	12	12	12	13	14	14	15
a_{n+30}	16	16	16	16	16	16	17	18	18	19

Column nonzero values

n	1	2	3	4	5	6	7	8	9	10	11
$t_{n,2}$	1	2	4	12	32	104	328	1 080	3 648	12 544	43 600
$t_{n,4}$	0	0	1	2	10	24	92	308	1 028	3 584	12 736
$t_{n,6}$	0	0	0	0	0	4	8	40	176	584	2 144
$t_{n,8}$	0	0	0	0	0	0	1	2	10	84	282
$t_{n,10}$	0	0	0	0	0	0	0	0	0	0	24

- Define $$a_n = \max\{k\colon\, t_{n,2k} > 0\}$$

n	1	2	3	4	5	6	7	8	9	10
a_n	1	1	2	2	2	3	4	4	4	4
a_{n+10}	5	6	6	7	8	8	8	8	8	9
a_{n+20}	10	10	11	12	12	12	13	14	14	15
a_{n+30}	16	16	16	16	16	16	17	18	18	19

- Lemma: The sequence (a_n) satisfies

$$a_1 = 1, \qquad a_n = \max\{k\colon\, k \leqslant 2a_{n-k}\}$$

Sequence of repeating elements of (a_n)

n	1	2	3	4	5	6	7	8	9	10
a_n	1	1	2	2	2	3	4	4	4	4
a_{n+10}	5	6	6	7	8	8	8	8	8	9
a_{n+20}	10	10	11	12	12	12	13	14	14	15
a_{n+30}	16	16	16	16	16	16	17	18	18	19

Define

$$\ell_n = \#\{k \colon a_k = n\}$$

Sequence of repeating elements of (a_n)

n	1	2	3	4	5	6	7	8	9	10
a_n	1	1	2	2	2	3	4	4	4	4
a_{n+10}	5	6	6	7	8	8	8	8	8	9
a_{n+20}	10	10	11	12	12	12	13	14	14	15
a_{n+30}	16	16	16	16	16	16	17	18	18	19

Define

$$\ell_n = \#\{k \colon a_k = n\}$$

We have

$$(\ell_n) = 2,\ 3,\ 1,\ 4,\ 1,\ 2,\ 1,\ 5,\ 1,\ 2,\ 1,\ 3,\ 1,\ 2,\ 1,\ 6,\ 1,\ 2,\ 1,\ \ldots$$

Description of (ℓ_n)

Proposition

The sequence (ℓ_n) satisfies

$$\ell_n = \begin{cases} p+2 & \text{if } n = 2^p, \\ p+1 & \text{if } n = 2^p a,\ a \text{ is odd},\ a > 1\,. \end{cases}$$

In particular,

- $\ell_{2n} = \ell_n + 1$ for even indices,
- $\ell_{2n+1} = 1$ for odd indices greater than 1,
- $\ell_1 = 2$.

Induction based on $a_n = \max\{k \colon k \leqslant 2a_{n-k}\}$

Description of (a_n)

Proposition

The sequence (a_n) satisfies

$$a_n = a_{n-1-a_{n-1}} + a_{n-2-a_{n-2}}, \qquad a_0 = a_1 = a_2 = 1$$

- Question. How to explain this recurrence combinatorially?

Description of (a_n)

Proposition

The sequence (a_n) satisfies

$$a_n = a_{n-1-a_{n-1}} + a_{n-2-a_{n-2}}, \qquad a_0 = a_1 = a_2 = 1$$

- Question. How to explain this recurrence combinatorially?
- (a_n) is known as a meta-Fibonacci sequence

Corollary (Tanny, 1992)

- $\displaystyle\lim_{n\to\infty} \frac{a_n}{n} = \frac{1}{2}$
- $\displaystyle\sum_{n=0}^{\infty} a_n z^n = z \sum_{n=0}^{\infty} \prod_{i=1}^{n} (z + z^{2^i})$

Trees of fixed height

$t_{n,2k,h} = \#\{\text{active trees with } n \text{ internal nodes } 2k \text{ anchors of height } h\}$

k								
4	0	0	0	0	0	0	1	0
3	0	0	0	0	0	4	0	0
2	0	0	0	2	6	0	0	0
1	0	0	4	4	0	0	0	0
$h = 3$	1	2	3	4	5	6	7	8

n

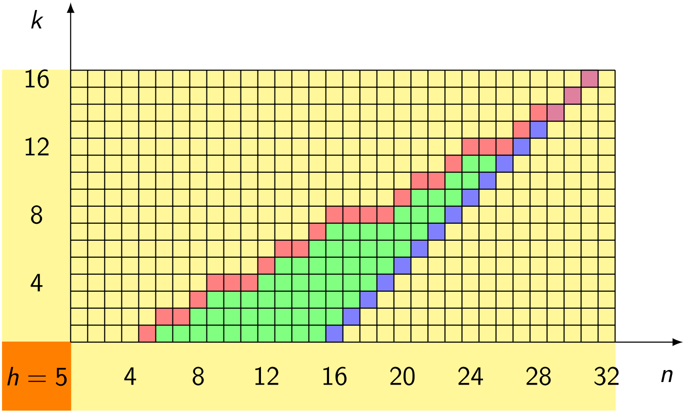
Trees of fixed height

$$t_{n,2k,h} = \#\{\text{active trees with } n \text{ internal nodes } 2k \text{ anchors of height } h\}$$

k																
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0
6	0	0	0	0	0	0	0	0	0	0	0	4	28	0	0	0
	0	0	0	0	0	0	0	0	0	0	24	56	0	0	0	0
4	0	0	0	0	0	0	0	2	6	60	70	0	0	0	0	0
	0	0	0	0	0	0	8	24	80	56	0	0	0	0	0	0
2	0	0	0	0	4	16	36	60	28	0	0	0	0	0	0	0
	0	0	0	8	16	24	24	8	0	0	0	0	0	0	0	0
$h = 4$		2		4		6		8		10		12		14		16 n

Trees of fixed height

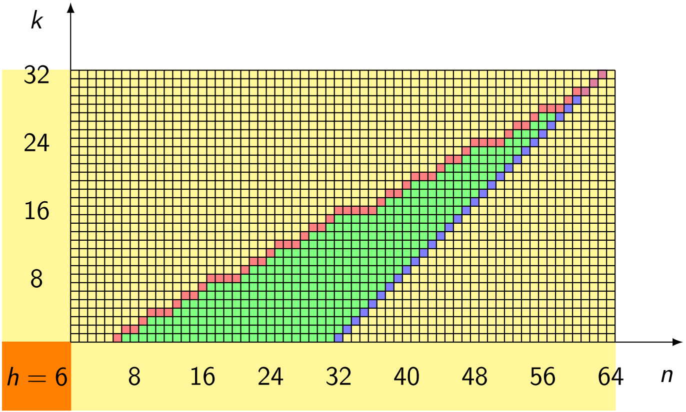
$t_{n,2k,h} = \#\{\text{active trees with } n \text{ internal nodes } 2k \text{ anchors of height } h\}$



$S_h = \{(n,k)\colon\, t_{n,2k,h} \neq 0\}, \qquad |S_h| = 2^{h-2}(2^{h-1} - h + 2)$

Trees of fixed height

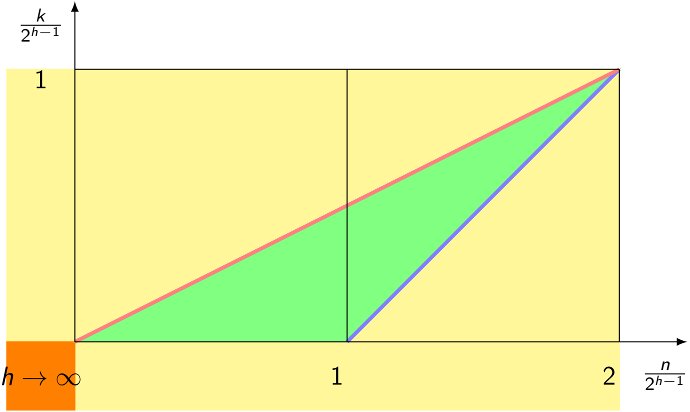
$t_{n,2k,h} = \#\{$active trees with n internal nodes $2k$ anchors of height $h\}$



$S_h = \{(n,k) \colon t_{n,2k,h} \neq 0\}$, $\qquad |S_h| = 2^{h-2}(2^{h-1} - h + 2)$

Trees of fixed height

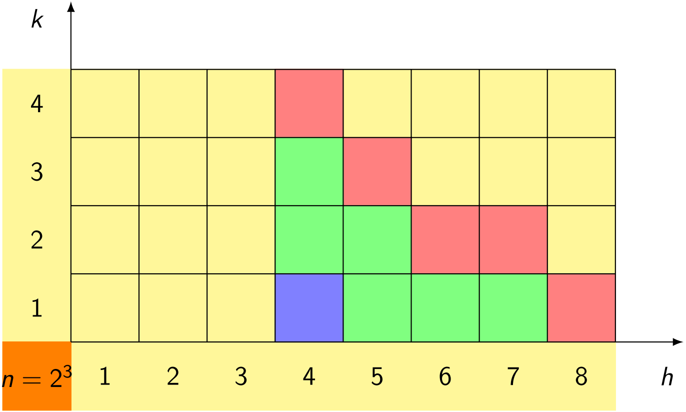
$t_{n,2k,h} = \#\{\text{active trees with } n \text{ internal nodes } 2k \text{ anchors of height } h\}$



$S_h = \{(n,k) \colon t_{n,2k,h} \neq 0\}$, $\quad |S_h| = 2^{h-2}(2^{h-1} - h + 2)$

Trees of fixed size

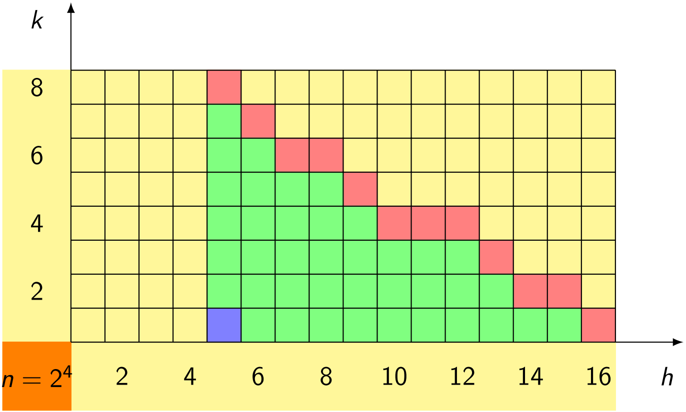
$t_{n,2k,h} = \#\{\text{active trees with } n \text{ internal nodes } 2k \text{ anchors of height } h\}$



$\widehat{S}_n = \{(h,k) \colon t_{n,2k,h} \neq 0\},$ $\qquad |\widehat{S}_{2^m}| = 2^{m-2}(2^m - m + 1)$

Trees of fixed size

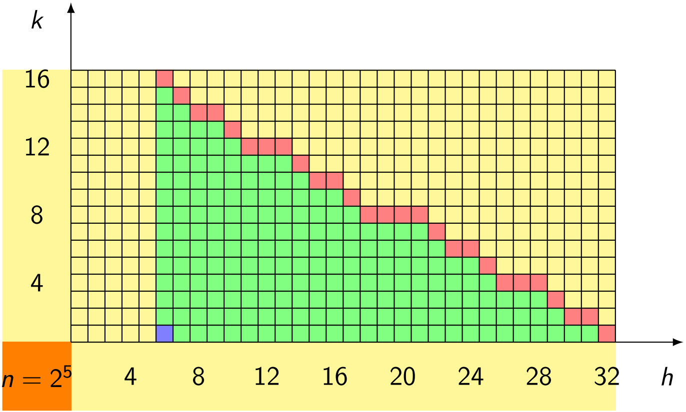
$t_{n,2k,h} = \#\{$active trees with n internal nodes $2k$ anchors of height $h\}$



$\widehat{S}_n = \{(h,k) \colon\, t_{n,2k,h} \neq 0\},$ $\qquad |\widehat{S}_{2^m}| = 2^{m-2}(2^m - m + 1)$

Trees of fixed size

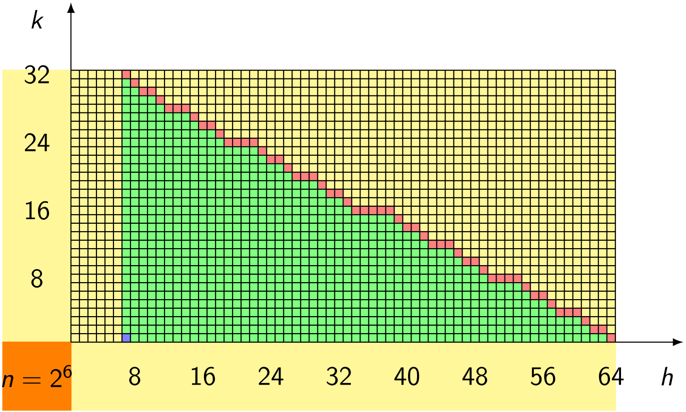
$$t_{n,2k,h} = \#\{\text{active trees with } n \text{ internal nodes } 2k \text{ anchors of height } h\}$$



$$\widehat{S}_n = \{(h,k) \colon t_{n,2k,h} \neq 0\}, \qquad |\widehat{S}_{2^m}| = 2^{m-2}(2^m - m + 1)$$

Trees of fixed size

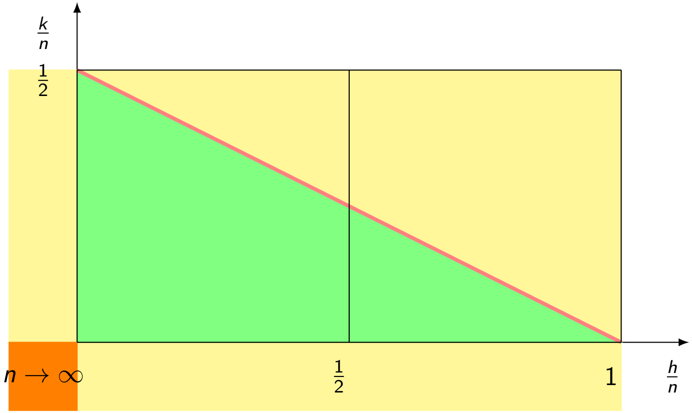
$t_{n,2k,h} = \#\{\text{active trees with } n \text{ internal nodes } 2k \text{ anchors of height } h\}$



$\widehat{S}_n = \{(h,k)\colon\ t_{n,2k,h} \neq 0\},$ $\qquad |\widehat{S}_{2^m}| = 2^{m-2}(2^m - m + 1)$

Trees of fixed size

$t_{n,2k,h} = \#\{$active trees with n internal nodes $2k$ anchors of height $h\}$



$\widehat{S}_n = \{(h,k)\colon\, t_{n,2k,h} \neq 0\}$, $\qquad |\widehat{S}_{2^m}| = 2^{m-2}(2^m - m + 1)$

Conclusion

1. Studied objects:
   - growing binary trees.
2. Related objects:
   - Mandelbrot polynomials,
   - meta-Fibonacci sequences.
3. Results:
   - relations for generating functions,
   - bounds for anchor distributions,
   - behavior of the maximal number of anchors,
   - limit forms of nonzero domains.

Thank you for your attention!

Literature I

Neil J. Calkin, Eunice Y. S. Chan, Robert M. Corless
Some Facts and Conjectures about Mandelbrot Polynomials
Maple Transactions, 2021.

Stephen M. Tanny
A well-behaved cousin of the Hofstadter sequence
Discrete Mathematics, 105 (1-3), 1992, pp. 227–239.