

Growing binary trees

Olivier Bodini¹, Antoine Genitrini², **Khaydar Nurligareev²**

(1) LAGA, USPN;

(2) LIP6, Sorbonne University

GASCom 2026

Hotel Bel Soggiorno, Malosco, Italy

June 11, 2026

Settings

Vertex types

- internal node
- anchor (active leaf)
- leaf (dead leaf)

Settings

Vertex types

- internal node $t = 0$ ○
- anchor (active leaf)
- leaf (dead leaf)

Growing process

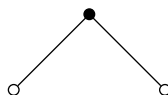
- At the beginning ($t = 0$),
our tree is an anchor ○

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 $t \geq 1$, replace each anchor ○ by
 - a leaf □
 - or a subtree 

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

Settings

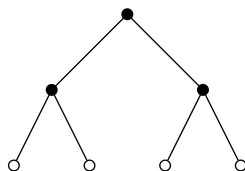
Vertex types

- internal node
- anchor (active leaf)
- leaf (dead leaf)

Growing process

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

$t = 2$



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

Settings

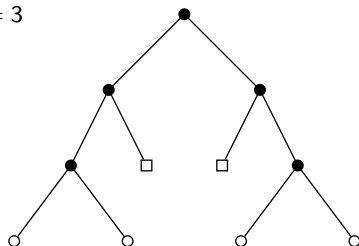
Vertex types

- internal node
- anchor (active leaf)
- leaf (dead leaf)

Growing process

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

$t = 3$



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

Settings

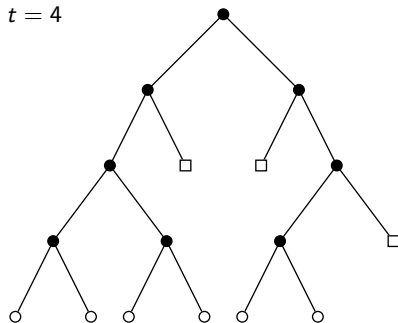
Vertex types

- internal node
- anchor (active leaf)
- leaf (dead leaf)

Growing process

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

$t = 4$



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

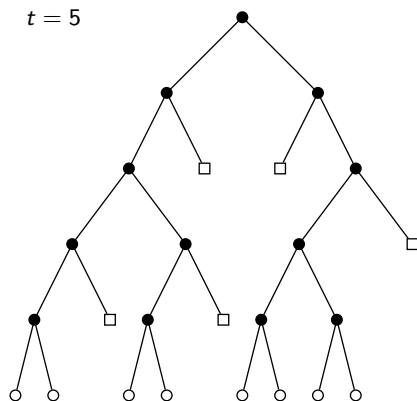
Settings

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 $t \geq 1$, replace each anchor ○ by
 - a leaf □
 - or a subtree



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

Counting

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

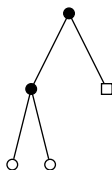
x



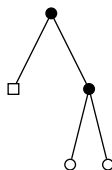
$x^2 z$



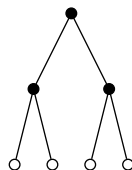
$x^2 z^2$



$x^2 z^2$



$x^4 z^3$



$$t_{0,1} = 1$$

$$t_{1,2} = 1$$

$$t_{2,2} = 2$$

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

Marking variables:

■ x marks anchors

■ z marks internal nodes

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

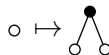
$$T(x, z) = x + x^2 z + 2x^2 z^2 + \dots$$

Equation

- Growing process replacement:

$$\circ \mapsto \square$$

$$x \mapsto 1$$



$$x \mapsto zx^2$$

- Equation:

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

Equation

- Growing process replacement:

$$\begin{array}{ll} \circ \mapsto \square & \circ \mapsto \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \circ \quad \circ \end{array} \\ x \mapsto 1 & x \mapsto zx^2 \end{array}$$

- Equation:

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

- $C(z) = 1 + zC^2(z)$

- C_n are Catalan numbers

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

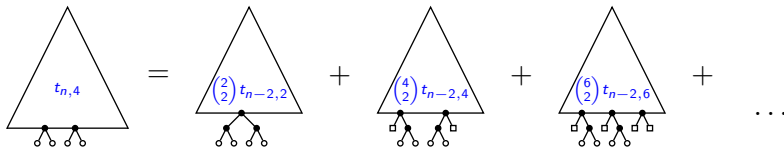
- Recurrent relation ($n, k > 0$):

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

Recurrent relations

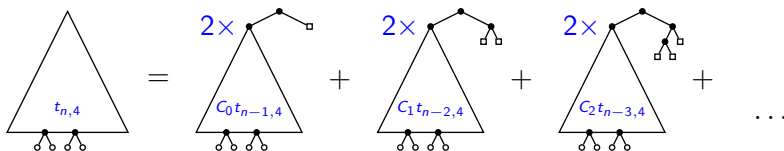
■ Exact:

$$t_{n,2k} = \sum_{m=\lceil k/2 \rceil}^{\infty} \binom{2m}{k} t_{n-k,2m} \quad (n > 0)$$



■ Asymptotic:

$$t_{n,2k} > 2 \sum_{m=0}^r C_m t_{n-m-1,2k} \quad (n \gg r)$$



Mandelbrot Polynomials

■ Define

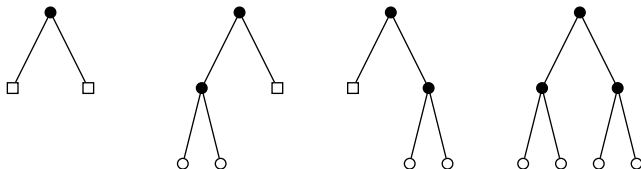
$$p_0(x, z) = x, \quad p_{n+1}(x, z) = 1 + zp_n^2(x, z)$$

■ $p_n(x, z)$ counts binary trees:

- of height at most n ,
- anchors are at level n ,
- leaves are at levels $k < n$

$$\text{for } k < n: \\ [z^k]p_n(x, z) = C_k$$

○



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

Mandelbrot Polynomials

- Define

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

- $p_n(x, z)$ counts binary trees:

- of height at most n ,
- anchors are at level n ,
- leaves are at levels $k < n$

$$\text{for } k < n: \\ [z^k]p_n(x, z) = C_k$$

- 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

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

- Equation:

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

- Relation:

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

- In terms of Mandelbrot polynomials:

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

Small values of $t_{n,2k}$

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

n	11	12	13	14	15	16
$t_{n,2}$	43 600	153 504	546 272	1 960 368	7 085 456	25 773 296
$t_{n,4}$	12 736	45 160	161 152	581 632	2 114 504	7 727 656
$t_{n,6}$	2 144	8 192	30 720	112 496	416 528	1 553 776
$t_{n,8}$	282	1 048	4 368	18 224	69 676	265 220
$t_{n,10}$	24	104	352	1 616	8 208	34 704
$t_{n,12}$	0	4	36	96	456	2 936
$t_{n,14}$	0	0	0	8	16	80
$t_{n,16}$	0	0	0	0	1	2

Anchor distributions

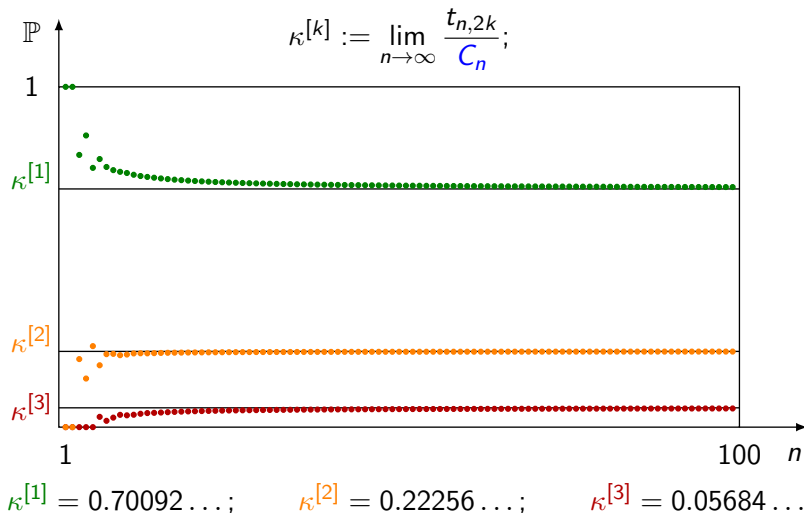
n	1	2	3	4	5	6	7	8	9	10	11
$t_{n,2}$	1	2	4	12	32	104	328	1080	3648	12544	43600
$t_{n,4}$	0	0	1	2	10	24	92	308	1028	3584	12736
$t_{n,6}$	0	0	0	0	0	4	8	40	176	584	2144
$t_{n,8}$	0	0	0	0	0	0	1	2	10	84	282
$t_{n,10}$	0	0	0	0	0	0	0	0	0	0	24
$t_{n,12}$	0	0	0	0	0	0	0	0	0	0	0
C_n	1	2	5	14	42	132	429	1430	4862	16796	58786

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



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	1080	3648	12544	43600
$t_{n,4}$	0	0	1	2	10	24	92	308	1028	3584	12736
$t_{n,6}$	0	0	0	0	0	4	8	40	176	584	2144
$t_{n,8}$	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 $a_n = \max\{k : t_{n,2k} > 0\}$

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

$$a_1 = 1, \quad a_n = \max\{k : k \leq 2a_{n-k}\}$$

One property of (a_n)

Lemma

The sequence (a_n) satisfies

$$a_1 = 1, \quad a_n = \max\{k : k \leq 2a_{n-k}\}$$

We have

$$t_{n,2k} = \sum_{m=\lceil k/2 \rceil}^{a_{n-k}} \binom{2m}{k} t_{n-k,2m}$$

and

$$t_{n,2k} \neq 0 \quad \Leftrightarrow \quad \exists m : k \leq 2m \leq 2a_{n-k}$$

Hence,

$$k \leq 2a_{n-k}$$

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

$$b_n = \#\{k: a_k = n\}$$

We have

$$(b_n) = 2, 3, 1, 4, 1, 2, 1, 5, 1, \textcolor{blue}{2}, 1, 3, 1, 2, 1, 6, 1, 2, 1, \dots$$

Description of (b_n)

Proposition

The sequence (b_n) satisfies

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

In particular,

- $b_{2n} = b_n + 1$ for even indices,
- $b_{2n+1} = 1$ for odd indices greater than 1,
- $b_1 = 2$.

Induction based on $a_n = \max\{k : k \leq 2a_{n-k}\}$

Description of (a_n)

Proposition

The sequence (a_n) satisfies

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

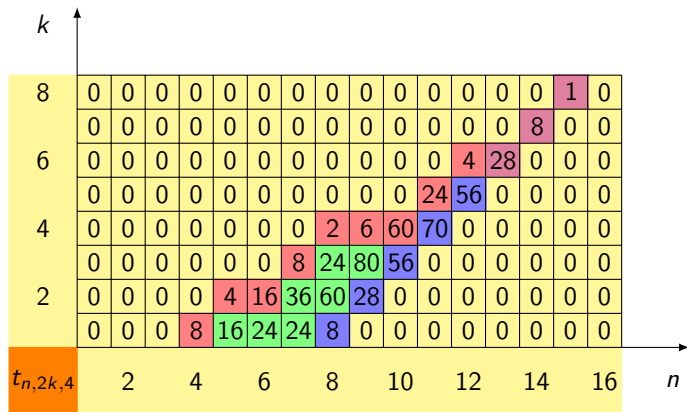
- Question. How to explain this recurrence combinatorially?
- (a_n) is known as a **meta-Fibonacci sequence**

Corollary (Tanny, 1992)

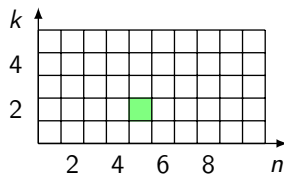
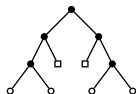
- $\lim_{n \rightarrow \infty} \frac{a_n}{n} = \frac{1}{2}$
- $\sum_{n=0}^{\infty} a_n z^n = z \sum_{n=0}^{\infty} \prod_{i=1}^n (z + z^{2^i})$

Trees of fixed height

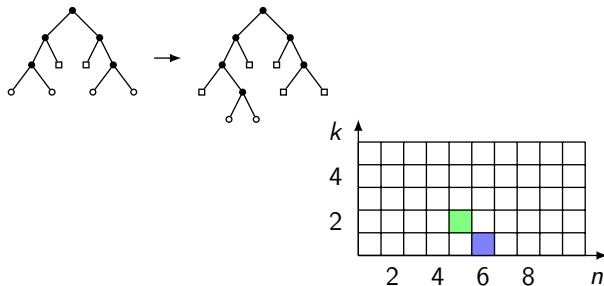
$$t_{n,2k,h} = \#\{\text{active trees with } n \text{ internal nodes } 2k \text{ anchors of height } h\}$$



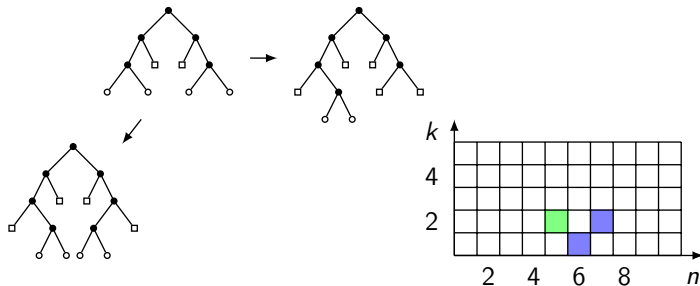
From h to $h + 1$



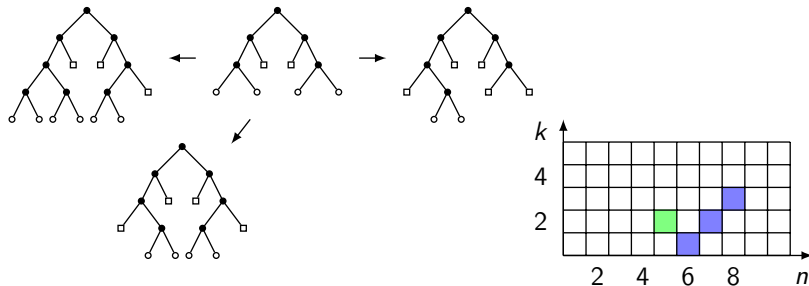
From h to $h + 1$



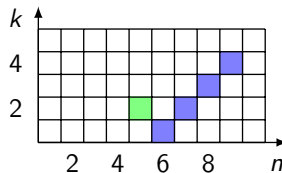
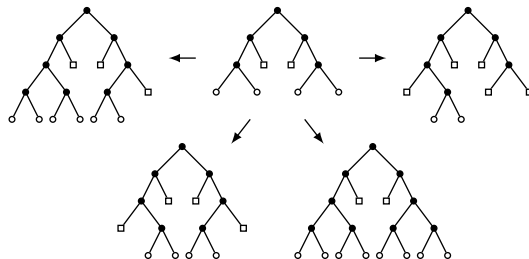
From h to $h + 1$



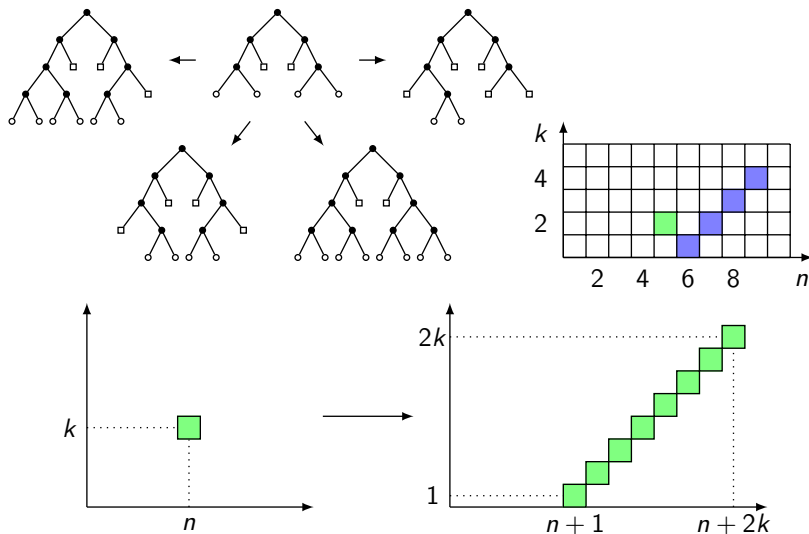
From h to $h + 1$



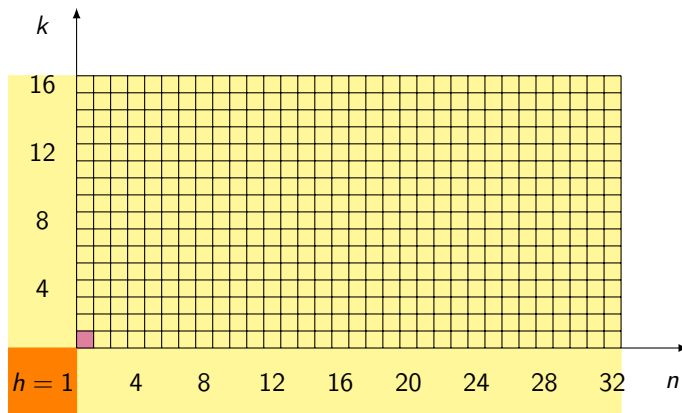
From h to $h + 1$



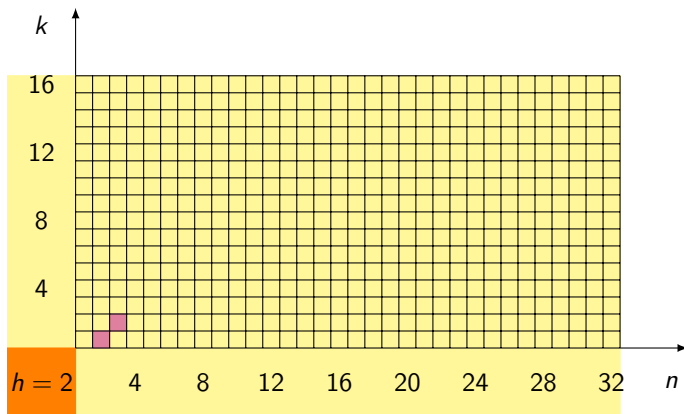
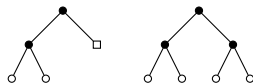
From h to $h + 1$



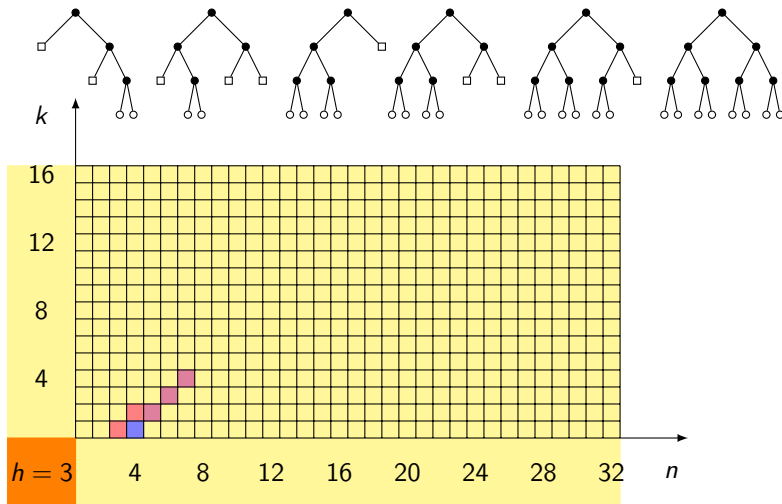
Non-zero domain properties



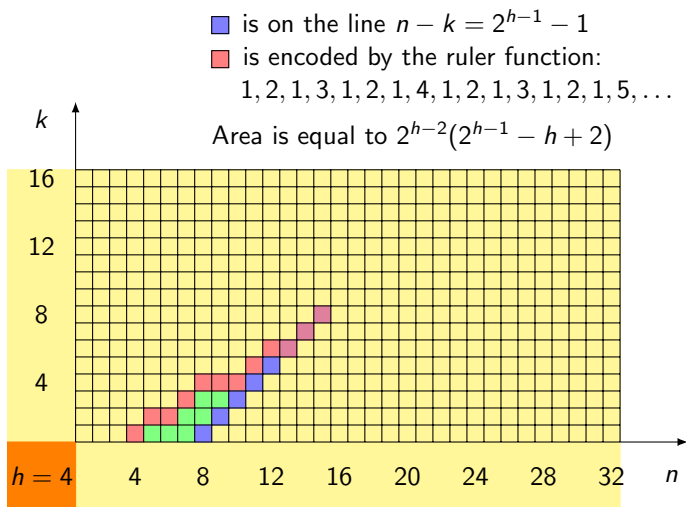
Non-zero domain properties



Non-zero domain properties



Non-zero domain properties



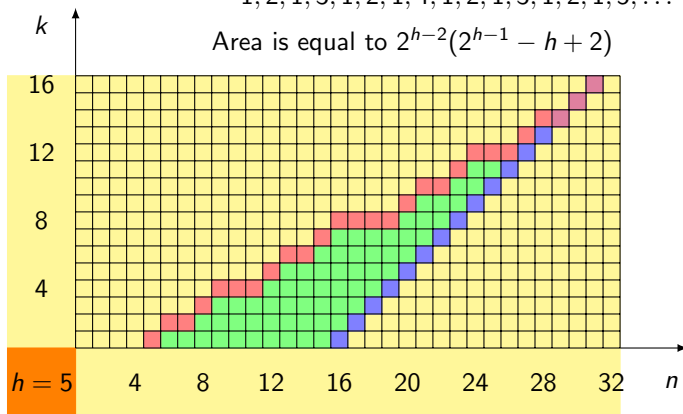
Non-zero domain properties

■ is on the line $n - k = 2^{h-1} - 1$

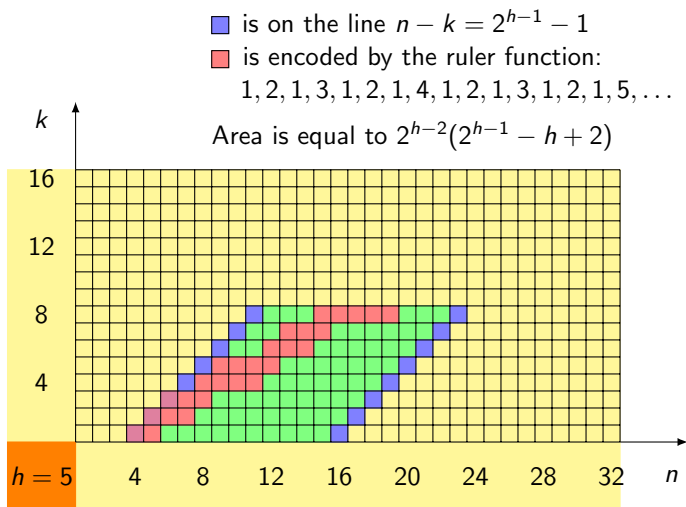
■ is encoded by the ruler function:

1, 2, 1, 3, 1, 2, 1, 4, 1, 2, 1, 3, 1, 2, 1, 5, ...

Area is equal to $2^{h-2}(2^{h-1} - h + 2)$

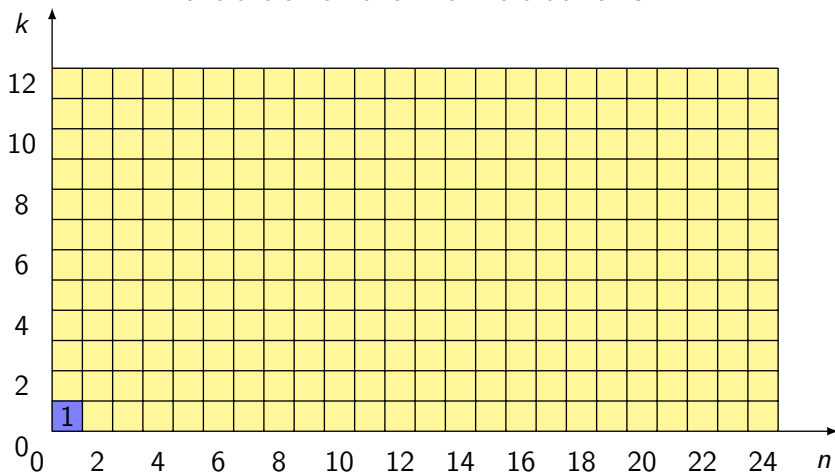


Non-zero domain properties



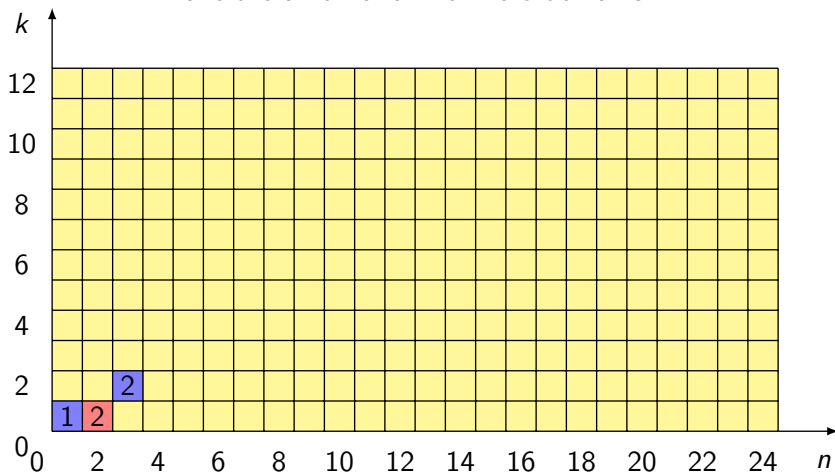
Behavior of the sequence $a_n = \max\{k: t_{n,2k} > 0\}$

Take the union of all non-zero domains



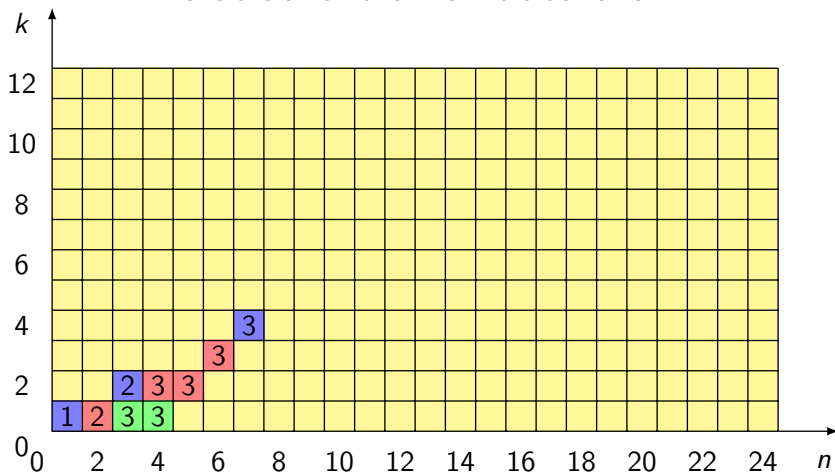
Behavior of the sequence $a_n = \max\{k: t_{n,2k} > 0\}$

Take the union of all non-zero domains



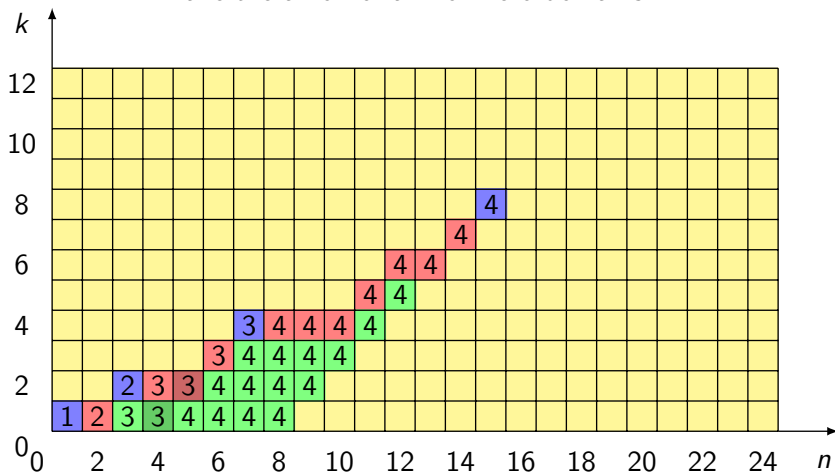
Behavior of the sequence $a_n = \max\{k: t_{n,2k} > 0\}$

Take the union of all non-zero domains



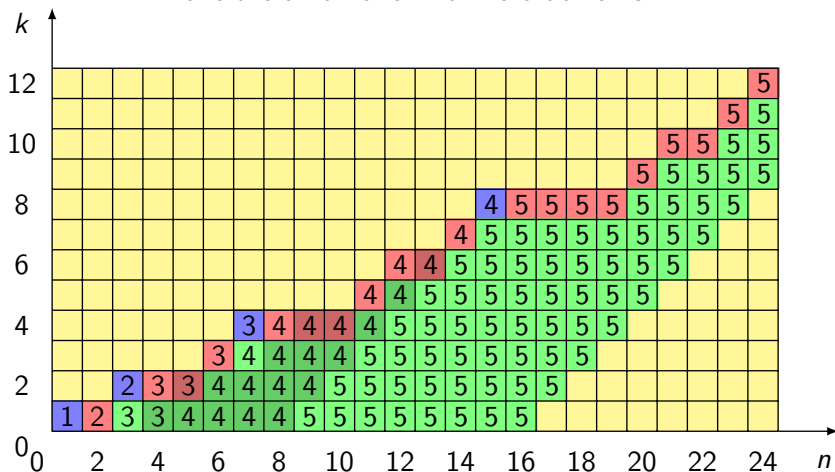
Behavior of the sequence $a_n = \max\{k: t_{n,2k} > 0\}$

Take the union of all non-zero domains



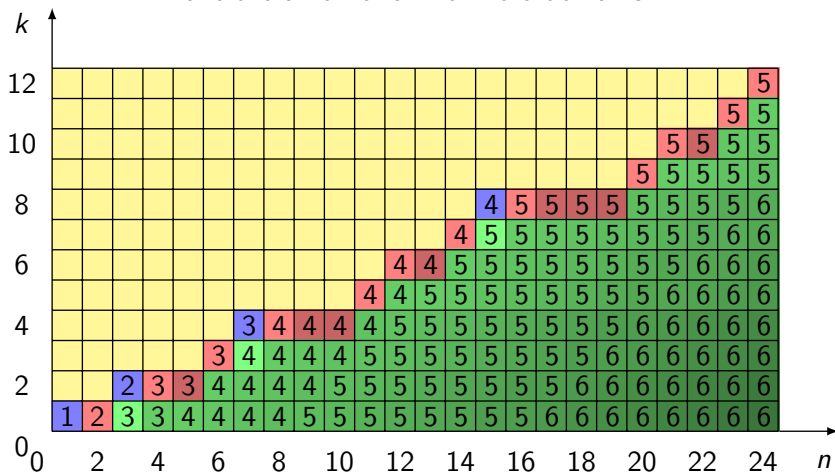
Behavior of the sequence $a_n = \max\{k: t_{n,2k} > 0\}$

Take the union of all non-zero domains



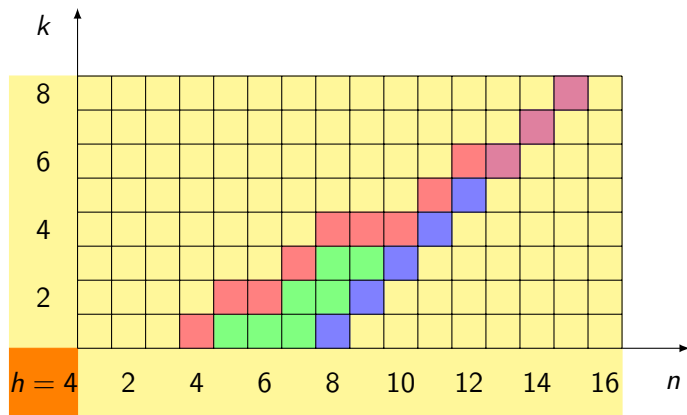
Behavior of the sequence $a_n = \max\{k: t_{n,2k} > 0\}$

Take the union of all non-zero domains



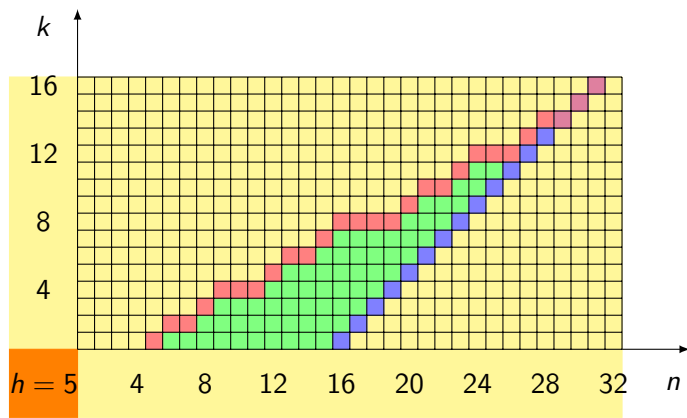
Scaling limit

The scaling limit of the non-zero domain (normalized by 2^{h-1}), as $h \rightarrow \infty$, is ...



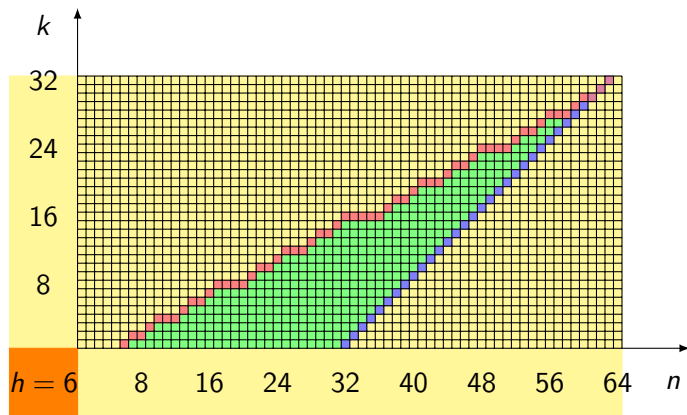
Scaling limit

The scaling limit of the non-zero domain (normalized by 2^{h-1}), as $h \rightarrow \infty$, is ...



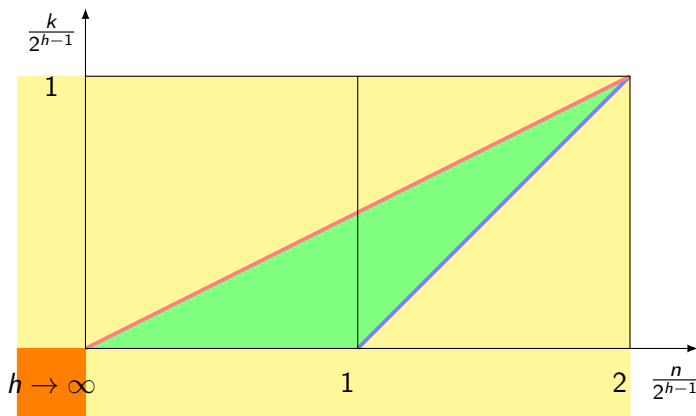
Scaling limit

The scaling limit of the non-zero domain (normalized by 2^{h-1}), as $h \rightarrow \infty$, is ...

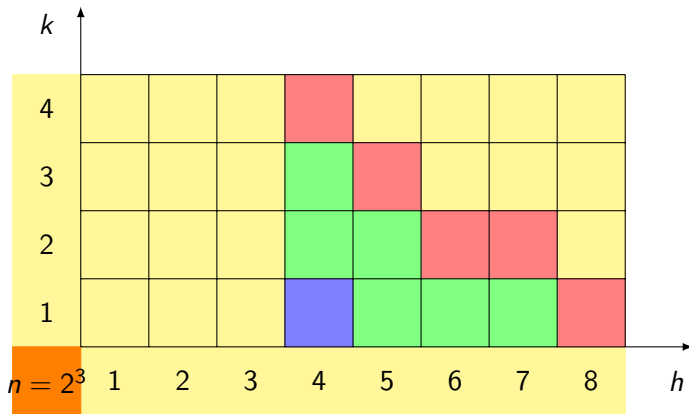


Scaling limit

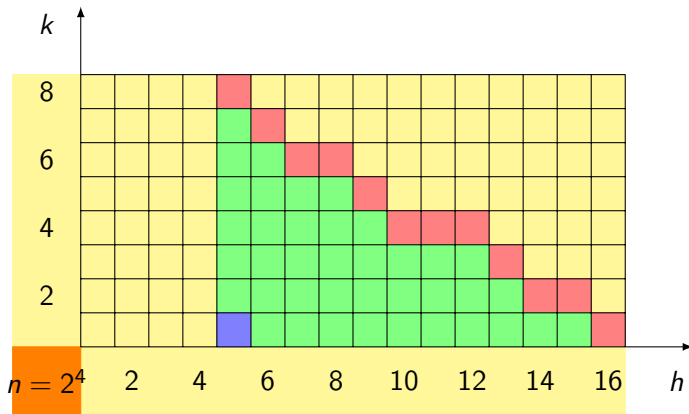
The scaling limit of the non-zero domain (normalized by 2^{h-1}), as $h \rightarrow \infty$, is the triangle with vertices $(0,0)$, $(1,0)$, and $(2,1)$.



Trees of fixed size



Trees of fixed size



Trees of fixed size

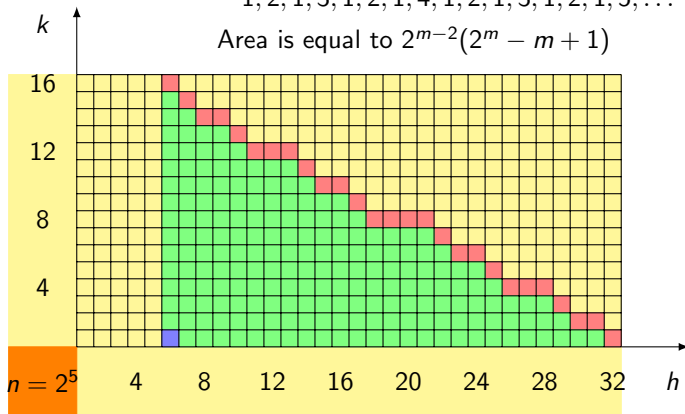
For $n = 2^m$:

■ is at $(m+1, 1)$

■ is encoded by the ruler function:

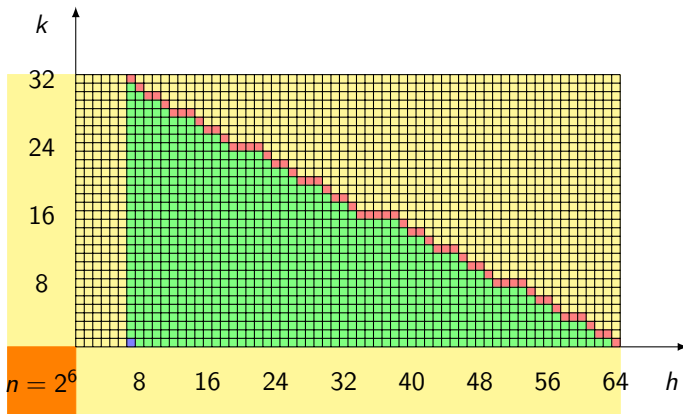
1, 2, 1, 3, 1, 2, 1, 4, 1, 2, 1, 3, 1, 2, 1, 5, ...

Area is equal to $2^{m-2}(2^m - m + 1)$



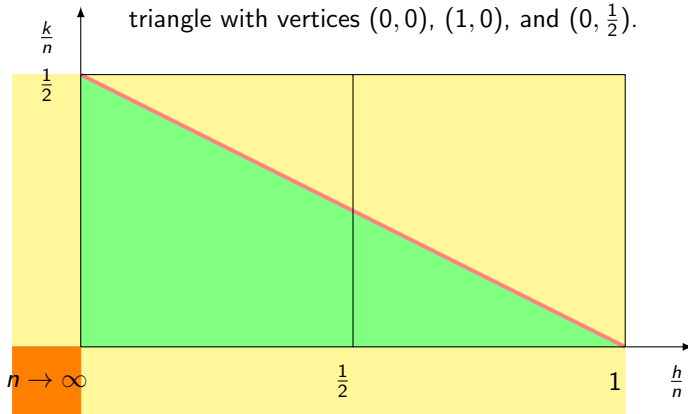
Trees of fixed size

The scaling limit of the non-zero domain
(normalized by n), as $n \rightarrow \infty$, is ...



Trees of fixed size

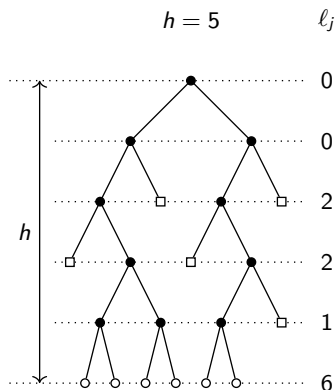
The scaling limit of the non-zero domain (normalized by n), as $n \rightarrow \infty$, is the triangle with vertices $(0,0)$, $(1,0)$, and $(0, \frac{1}{2})$.



Profile of a tree

- h = height
- $\ell_j = \#\{\text{leaves at } j\text{th level}\}$
- $(\ell_0, \dots, \ell_h) = \text{tree profile}$

$$\sum_{j=0}^h \frac{\ell_j}{2^j} = 1$$



$$\frac{2}{2^2} + \frac{2}{2^3} + \frac{1}{2^4} + \frac{6}{2^5} = \frac{1}{2} + \frac{1}{4} + \frac{1}{16} + \frac{3}{16} = 1$$

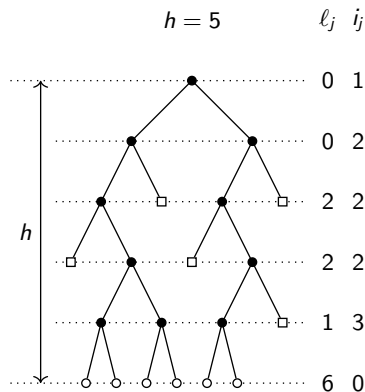
Profile of a tree

- $h = \text{height}$
- $\ell_j = \#\{\text{leaves at } j\text{th level}\}$
- $(\ell_0, \dots, \ell_h) = \text{tree profile}$

$$\sum_{j=0}^h \frac{\ell_j}{2^j} = 1$$

- $i_j = \#\{\text{nodes at } j\text{th level}\}$

$$i_0 = 1, \quad i_j + \ell_j = 2i_{j-1}$$



$$\# \left\{ \begin{array}{l} \text{trees with a valid} \\ \text{profile } (\ell_0, \dots, \ell_h) \end{array} \right\} = \prod_{j=0}^{h-1} \binom{2i_j}{\ell_{j+1}}$$

Sampling trees with a given profile

Input: $L = (\ell_0, \dots, \ell_h)$

Output: uniform tree with profile L

1: **function** UNIFORM_TREE(L)

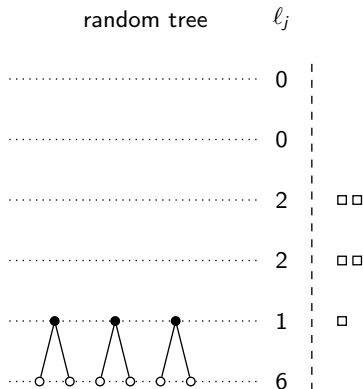
2: T = sequence of $\ell_h/2$ nodes,
 each one pointing to 2 leaves

3: **for** i from $h-1$ to 0 **do**

4: T' = shuffling T with ℓ_i leaves

5: T = sequence of $|T'|/2$ nodes,
 with the i th node pointing
 to nodes $2i$ and $2i+1$ of T'

6: **return** the first node of T



The algorithm is optimal in time, space, and random bit consumption.

Sampling trees with a given profile

Input: $L = (\ell_0, \dots, \ell_h)$

Output: uniform tree with profile L

1: **function** UNIFORM_TREE(L)

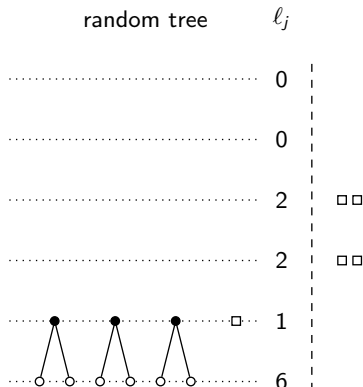
2: T = sequence of $\ell_h/2$ nodes,
 each one pointing to 2 leaves

3: **for** i from $h-1$ to 0 **do**

4: T' = shuffling T with ℓ_i leaves

5: T = sequence of $|T'|/2$ nodes,
 with the i th node pointing
 to nodes $2i$ and $2i+1$ of T'

6: **return** the first node of T



The algorithm is optimal in time, space, and random bit consumption.

Sampling trees with a given profile

Input: $L = (\ell_0, \dots, \ell_h)$

Output: uniform tree with profile L

1: **function** UNIFORM_TREE(L)

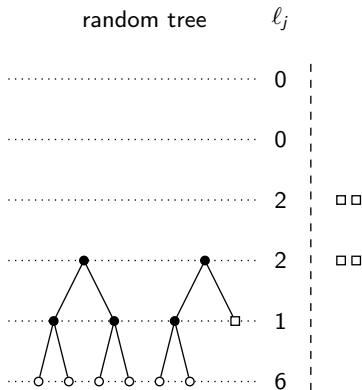
2: T = sequence of $\ell_h/2$ nodes,
each one pointing to 2 leaves

3: **for** i from $h-1$ to 0 **do**

4: T' = shuffling T with ℓ_i leaves

5: T = sequence of $|T'|/2$ nodes,
with the i th node pointing
to nodes $2i$ and $2i+1$ of T'

6: **return** the first node of T



The algorithm is optimal in time, space, and random bit consumption.

Sampling trees with a given profile

Input: $L = (\ell_0, \dots, \ell_h)$

Output: uniform tree with profile L

1: **function** UNIFORM_TREE(L)

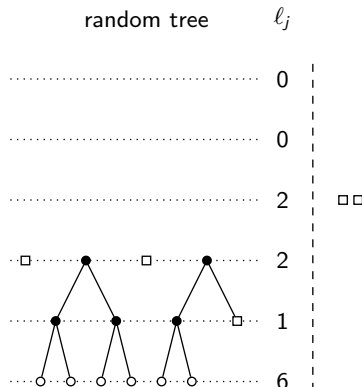
2: T = sequence of $\ell_h/2$ nodes,
 each one pointing to 2 leaves

3: **for** i from $h-1$ to 0 **do**

4: $T' =$ shuffling T with ℓ_i leaves

5: $T =$ sequence of $|T'|/2$ nodes,
 with the i th node pointing
 to nodes $2i$ and $2i+1$ of T'

6: **return** the first node of T



The algorithm is optimal in time, space, and random bit consumption.

Sampling trees with a given profile

Input: $L = (\ell_0, \dots, \ell_h)$

Output: uniform tree with profile L

1: **function** UNIFORM_TREE(L)

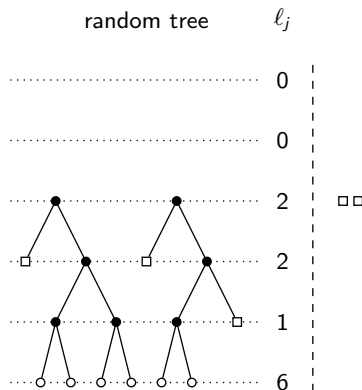
2: T = sequence of $\ell_h/2$ nodes,
each one pointing to 2 leaves

3: **for** i from $h-1$ to 0 **do**

4: T' = shuffling T with ℓ_i leaves

5: T = sequence of $|T'|/2$ nodes,
with the i th node pointing
to nodes $2i$ and $2i+1$ of T'

6: **return** the first node of T



The algorithm is optimal in time, space, and random bit consumption.

Sampling trees with a given profile

Input: $L = (\ell_0, \dots, \ell_h)$

Output: uniform tree with profile L

1: **function** UNIFORM_TREE(L)

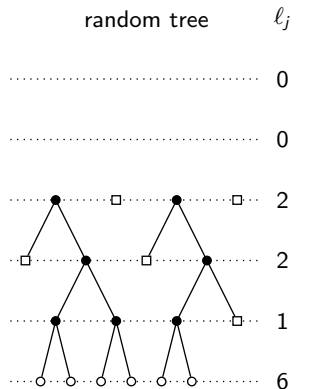
2: T = sequence of $\ell_h/2$ nodes,
each one pointing to 2 leaves

3: **for** i from $h-1$ to 0 **do**

4: T' = shuffling T with ℓ_i leaves

5: T = sequence of $|T'|/2$ nodes,
with the i th node pointing
to nodes $2i$ and $2i+1$ of T'

6: **return** the first node of T



The algorithm is optimal in time, space, and random bit consumption.

Sampling trees with a given profile

Input: $L = (\ell_0, \dots, \ell_h)$

Output: uniform tree with profile L

1: **function** UNIFORM_TREE(L)

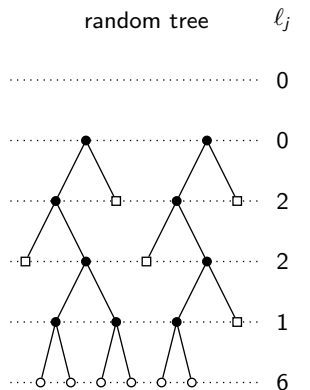
2: T = sequence of $\ell_h/2$ nodes,
each one pointing to 2 leaves

3: **for** i from $h-1$ to 0 **do**

4: T' = shuffling T with ℓ_i leaves

5: T = sequence of $|T'|/2$ nodes,
with the i th node pointing
to nodes $2i$ and $2i+1$ of T'

6: **return** the first node of T



The algorithm is optimal in time, space, and random bit consumption.

Sampling trees with a given profile

Input: $L = (\ell_0, \dots, \ell_h)$

Output: uniform tree with profile L

1: **function** UNIFORM_TREE(L)

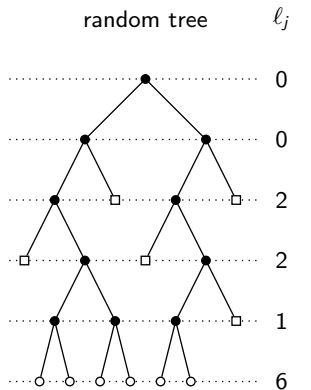
2: T = sequence of $\ell_h/2$ nodes,
each one pointing to 2 leaves

3: **for** i from $h-1$ to 0 **do**

4: $T' =$ shuffling T with ℓ_i leaves

5: $T =$ sequence of $|T'|/2$ nodes,
with the i th node pointing
to nodes $2i$ and $2i+1$ of T'

6: **return** the first node of T



The algorithm is optimal in time, space, and random bit consumption.

Conclusion

- 1 Studied objects:
 - growing binary trees.
- 2 Related objects:
 - Mandelbrot polynomials,
 - meta-Fibonacci sequences.
- 3 Results:
 - relations for generating functions,
 - behavior of the maximal number of anchors,
 - limit forms of non-zero domains,
 - uniform sampling of trees with a given profile.

Conclusion

- 1 Studied objects:
 - growing binary trees.
- 2 Related objects:
 - Mandelbrot polynomials,
 - meta-Fibonacci sequences.
- 3 Results:
 - relations for generating functions,
 - behavior of the maximal number of anchors,
 - limit forms of non-zero domains,
 - uniform sampling of trees with a given profile.

Thank you for your attention!

Literature I



P. Flajolet, R. Sedgewick,
Analytic Combinatorics.
Cambridge University Press, 2009.



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.