Growing binary trees

Antoine Genitrini and Khaydar Nurligareev

LIP6, Sorbonne University

Workshop PAnDAG 2025

May 16, 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

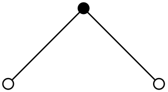
- At the beginning ($t = 0$), our tree is an anchor $\circ$

Settings

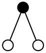
Vertex types

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

$t = 1$



Growing process

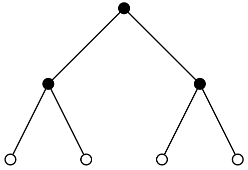
- At the beginning ($t = 0$), our tree is an anchor ○
- At any moment $t \geqslant 1$, replace each anchor ○ by
  - a leaf □
  - or a subtree 

Settings

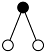
Vertex types

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

$t = 2$



Growing process

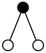
- At the beginning ($t = 0$), our tree is an anchor ○
- At any moment $t \geqslant 1$, replace each anchor ○ by
  - a leaf □
  - or a subtree 

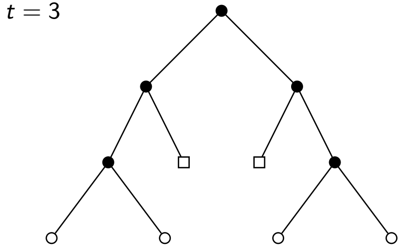
Settings

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 

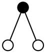


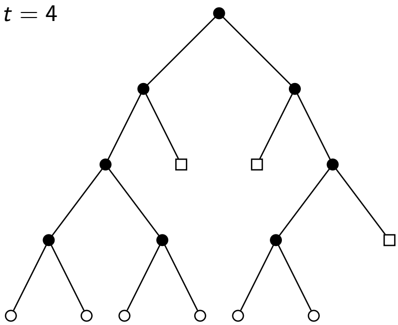
Settings

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 

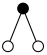


Settings

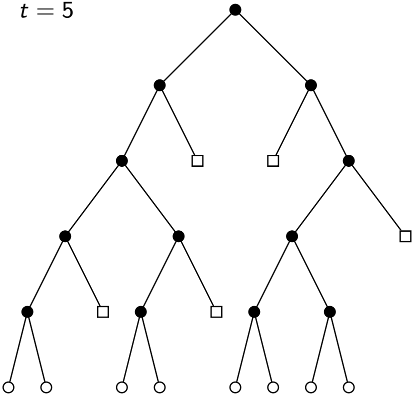
Vertex types

- ● 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 

$t = 5$

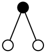


Settings

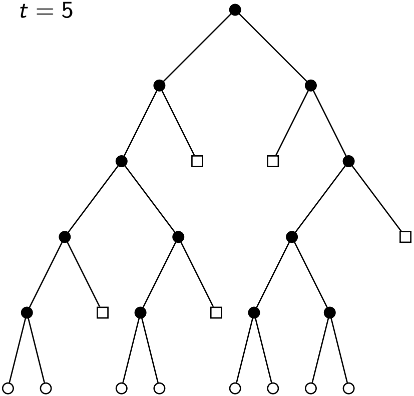
Vertex types

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

Growing process

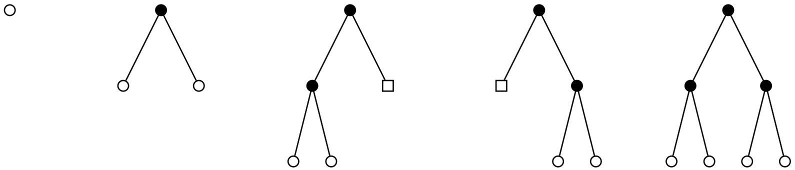
- At the beginning ($t = 0$), our tree is an anchor ○
- At any moment $t \geqslant 1$, replace each anchor ○ by
  - a leaf □
  - or a subtree 

$t = 5$



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

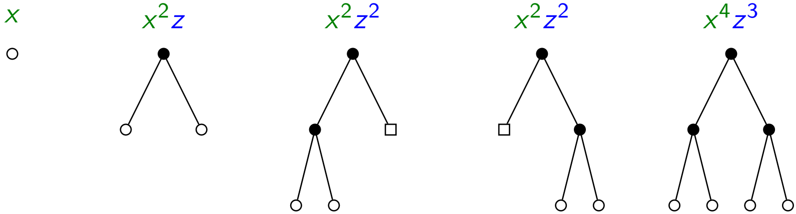
Some examples



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

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

Some examples



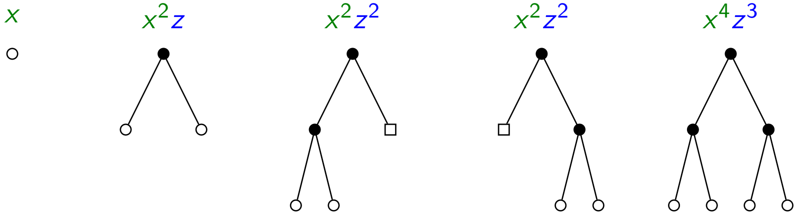
$$t_{n,m} = \{\text{active trees with } n \text{ internal nodes } m \text{ anchors}\}$$

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

Marking variables:

- x marks anchors
- z marks internal nodes

Some examples



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

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

Marking variables:

- x marks anchors
- z marks internal nodes

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

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

Generating function

- $t_{n,m} = \{$active trees with n internal nodes m anchors$\}$
- Generation function:

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

- Growing process replacement:

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

- Equation:

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

Some relations

- Equation (once again):

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

- Boundary conditions:

$$T(1,z) = C(z) \qquad \Leftrightarrow \qquad \sum_{m=1}^{\infty} t_{n,m} = C_n$$

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

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

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

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

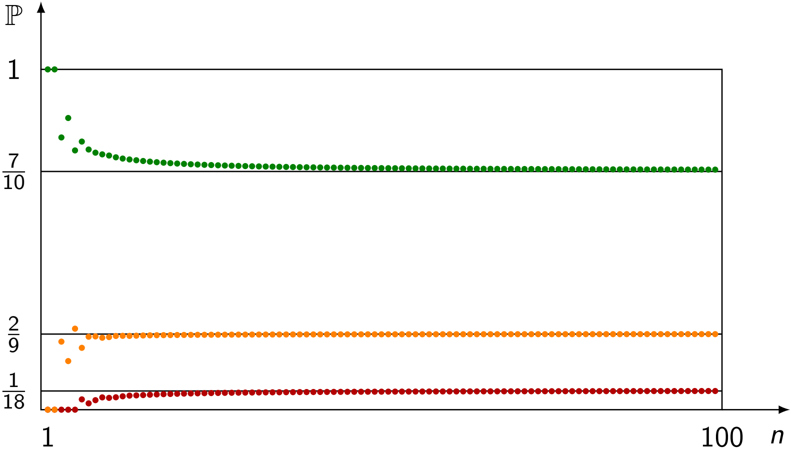
n	1	2	3	4	5	6	7	8	9	10	11
$t_{n,2}$	1	2	4	12	32	104	328	1 080	3 648	12 544	43 600
$t_{n,4}$	0	0	1	2	10	24	92	308	1 028	3 584	12 736
$t_{n,6}$	0	0	0	0	0	4	8	40	176	584	2 144
$t_{n,8}$	0	0	0	0	0	0	1	2	10	84	282
$t_{n,10}$	0	0	0	0	0	0	0	0	0	0	24
$t_{n,12}$	0	0	0	0	0	0	0	0	0	0	0
C_n	1	2	5	14	42	132	429	1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

It looks like

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

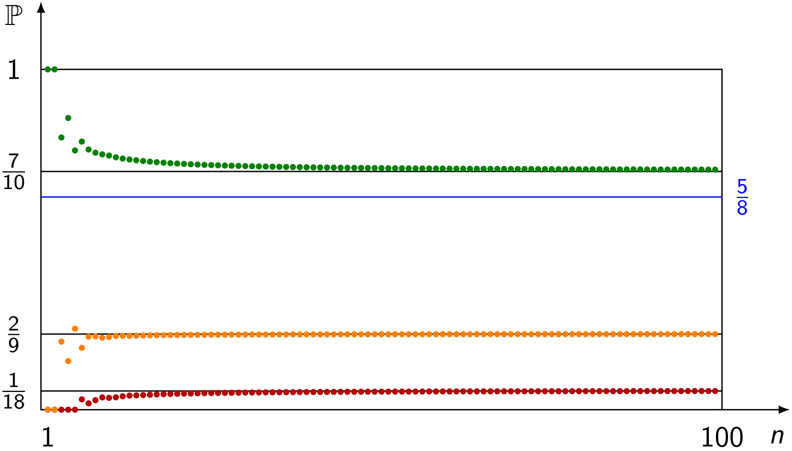
and there are some limits.

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



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

Simple lower bound

Proposition

The proportion of trees with two anchors satisfies

$$\liminf_{n\to\infty} \frac{t_{n,2}}{C_n} \geqslant \frac{1}{2}$$

Simple lower bound

Proposition

The proportion of trees with two anchors satisfies

$$\liminf_{n\to\infty} \frac{t_{n,2}}{C_n} \geqslant \frac{1}{2}$$

$$t_{n,2} \quad = \quad 2t_{n-1,2} + 4t_{n-1,4} + 6t_{n-1,6} + 8t_{n-1,8} + \ldots$$

Simple lower bound

Proposition

The proportion of trees with two anchors satisfies

$$\liminf_{n\to\infty} \frac{t_{n,2}}{C_n} \geqslant \frac{1}{2}$$

$$\begin{aligned} t_{n,2} &= 2t_{n-1,2} + 4t_{n-1,4} + 6t_{n-1,6} + 8t_{n-1,8} + \dots \\ &> 2t_{n-1,2} + 2t_{n-1,4} + 2t_{n-1,6} + 2t_{n-1,8} + \dots \end{aligned}$$

Simple lower bound

Proposition

The proportion of trees with two anchors satisfies

$$\liminf_{n\to\infty} \frac{t_{n,2}}{C_n} \geqslant \frac{1}{2}$$

$$\begin{aligned} t_{n,2} &= 2t_{n-1,2} + 4t_{n-1,4} + 6t_{n-1,6} + 8t_{n-1,8} + \dots \\ &> 2t_{n-1,2} + 2t_{n-1,4} + 2t_{n-1,6} + 2t_{n-1,8} + \dots \\ &= 2C_{n-1} \end{aligned}$$

Simple lower bound

Proposition

The proportion of trees with two anchors satisfies

$$\liminf_{n\to\infty} \frac{t_{n,2}}{C_n} \geqslant \frac{1}{2}$$

$$\begin{aligned} t_{n,2} &= 2t_{n-1,2} + 4t_{n-1,4} + 6t_{n-1,6} + 8t_{n-1,8} + \dots \\ &> 2t_{n-1,2} + 2t_{n-1,4} + 2t_{n-1,6} + 2t_{n-1,8} + \dots \\ &= 2C_{n-1} \end{aligned}$$

$$C_n \sim \frac{4^n}{n^{3/2}\sqrt{\pi}} \qquad \Rightarrow \qquad \frac{t_{n,2}}{C_n} > \frac{2C_{n-1}}{C_n} \sim \frac{1}{2}$$

Elaborated lower bound

Proposition

The proportion of trees with two anchors satisfies

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

Elaborated lower bound

Proposition

The proportion of trees with two anchors satisfies

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

$$t_{n+1,2} = 3C_n + \big(-t_{n,2} + t_{n,4} + 3t_{n,6} + 5t_{n,8} + \dots \big)$$

Elaborated lower bound

Proposition

The proportion of trees with two anchors satisfies

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

$$t_{n+1,2} = 3C_n + \big(- t_{n,2} + t_{n,4} + 3t_{n,6} + 5t_{n,8} + \dots \big)$$

$$t_{n,2} = 2C_{n-1} + \big(2t_{n-1,4} + 4t_{n-1,6} + 6t_{n-1,8} + \dots \big)$$

Elaborated lower bound

Proposition

The proportion of trees with two anchors satisfies

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

$$t_{n+1,2} = 3C_n + \big(- t_{n,2} + t_{n,4} + 3t_{n,6} + 5t_{n,8} + \dots \big)$$

$$t_{n,2} = 2C_{n-1} + \big(2t_{n-1,4} + 4t_{n-1,6} + 6t_{n-1,8} + \dots \big)$$

$$t_{n+1,2k} \geqslant 2t_{n,2k} \qquad \Rightarrow \qquad \frac{t_{n+1,2}}{C_{n+1}} > \frac{3C_n - 2C_{n-1}}{C_{n+1}} \sim \frac{5}{8}$$

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
$t_{n,12}$	0	0	0	0	0	0	0	0	0	0	0

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

One property of (a_n)

Proposition

The sequence (a_n) satisfies

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

One property of (a_n)

Proposition

The sequence (a_n) satisfies

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

We have

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

and

$$t_{n,2k} \neq 0 \qquad \Rightarrow \qquad \exists\, m \colon \; k \leqslant 2m \leqslant 2a_{n-k}$$

Hence,

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

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?

Corollary

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