On the number of variables in special classes of random lambda-terms

B. Gittenberger[1] I. Larcher[1]

[1]Institute of Discrete Mathematics and Geometry
Technische Universität Wien

CLA Paris, 2018

Lambda-terms

The set Λ of lambda-terms is defined by the following grammar:

- x variable $\Rightarrow x \in \Lambda$.
- $M, N \in \Lambda \Rightarrow MN \in \Lambda$. $\rightarrow$ *application*
- x variable and $M \in \Lambda \Rightarrow \lambda x.M \in \Lambda$. $\rightarrow$ *abstraction*

Lambda-terms

The set Λ of lambda-terms is defined by the following grammar:

- x variable $\Rightarrow x \in \Lambda$.
- $M, N \in \Lambda \Rightarrow MN \in \Lambda$. $\rightarrow$ *application*
- x variable and $M \in \Lambda \Rightarrow \lambda x.M \in \Lambda$. $\rightarrow$ *abstraction*

examples: $\lambda x.((\lambda y.yz)x)$, $(\lambda x.x)(\lambda y.zy)$, $\lambda x.(x(\lambda y.yx))$

Lambda-terms

The set Λ of lambda-terms is defined by the following grammar:

- x variable $\Rightarrow x \in \Lambda$.
- $M, N \in \Lambda \Rightarrow MN \in \Lambda$. $\rightarrow$ *application*
- x variable and $M \in \Lambda \Rightarrow \lambda x.M \in \Lambda$. $\rightarrow$ *abstraction*

examples: $\lambda x.((\lambda y.yz)x)$, $(\lambda x.x)(\lambda y.zy)$, $\lambda x.(x(\lambda y.yx))$

$\rightarrow$ Every λ binds exactly one variable (but maybe none or several occurrences of this variable).

Lambda-terms

The set Λ of lambda-terms is defined by the following grammar:

- x variable $\Rightarrow x \in \Lambda$.
- $M, N \in \Lambda \Rightarrow MN \in \Lambda$. $\rightarrow$ *application*
- x variable and $M \in \Lambda \Rightarrow \lambda x.M \in \Lambda$. $\rightarrow$ *abstraction*

examples: $\lambda x.((\lambda y.yz)x)$, $(\lambda x.x)(\lambda y.zy)$, $\lambda x.(x(\lambda y.yx))$

$\rightarrow$ Every λ binds exactly one variable (but maybe none or several occurrences of this variable).
$\rightarrow$ Every variable is bound by at most one λ.

Lambda-terms

The set Λ of lambda-terms is defined by the following grammar:

- x variable $\Rightarrow x \in \Lambda$.
- $M, N \in \Lambda \Rightarrow MN \in \Lambda$. $\rightarrow$ *application*
- x variable and $M \in \Lambda \Rightarrow \lambda x.M \in \Lambda$. $\rightarrow$ *abstraction*

examples: $\lambda x.((\lambda y.yz)x)$, $(\lambda x.x)(\lambda y.zy)$, $\lambda x.(x(\lambda y.yx))$

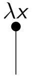
$\rightarrow$ Every λ binds exactly one variable (but maybe none or several occurrences of this variable).
$\rightarrow$ Every variable is bound by at most one λ.
$\rightarrow$ Variables that are not bound by any λ are called free.

Lambda-DAGs

$$\lambda x.((\lambda y.(xy))z)$$

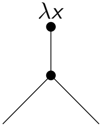
Lambda-DAGs

$\lambda x.((\lambda y.(xy))z)$



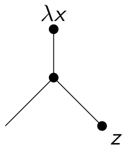
Lambda-DAGs

$\lambda x.((\lambda y.(xy))z)$



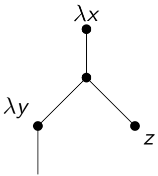
Lambda-DAGs

$\lambda x.((\lambda y.(xy))z)$



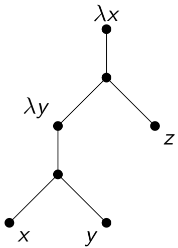
Lambda-DAGs

$\lambda x.((\lambda y.(xy))z)$



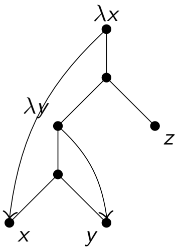
Lambda-DAGs

$\lambda x.((\lambda y.(xy))z)$



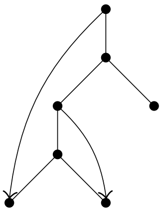
Lambda-DAGs

$\lambda x.((\lambda y.(xy))z)$



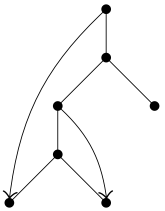
Lambda-DAGs

$\lambda x.((\lambda y.(xy))z)$



Lambda-DAGs

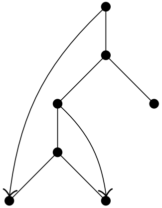
$\lambda x.((\lambda y.(xy))z)$



- α-equivalence

Lambda-DAGs

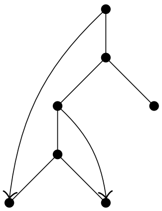
$\lambda x.((\lambda y.(xy))z)$



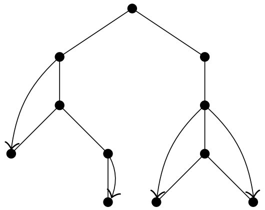
- α-equivalence
- contexts

Lambda-DAGs

$\lambda x.((\lambda y.(xy))z)$



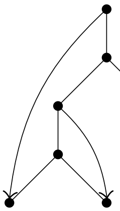
$(\lambda x.(x(\lambda y.y))(\lambda x.(\lambda y.yy))$



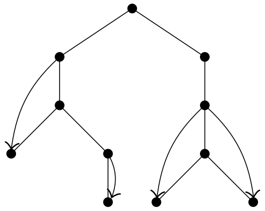
- α-equivalence
- contexts

Lambda-DAGs

$\lambda x.((\lambda y.(xy))z)$



$(\lambda x.(x(\lambda y.y))(\lambda x.(\lambda y.yy))$

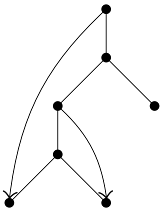


- α-equivalence
- contexts

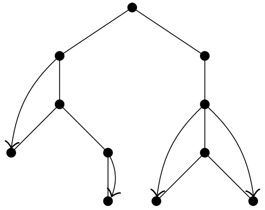
$|x| = 1 \qquad |MN| = 1 + |M| + |N| \qquad |\lambda x.M| = 1 + |M|$

Lambda-DAGs

$|\lambda x.((\lambda y.(xy))z)| = 7$



$|(\lambda x.(x(\lambda y.y))(\lambda x.(\lambda y.yy))| \;=\; 11$



- α-equivalence
- contexts

$|x| = 1$ $\qquad |MN| = 1 + |M| + |N|$ $\qquad |\lambda x.M| = 1 + |M|$

- **Generating functions:**

$a_n \ldots$ number of lambda-terms of size n

$A(z) = \sum_{n \geq 0} a_n z^n \ldots$ GF of $(a_n)_{n \geq 0}$

$\Rightarrow [z^n]A(z) = a_n$

- **Generating functions:**

$a_{n,k}$ ... number of lambda-terms of size n with k variables

$A(z,u) = \sum_{n,k\geq 0} a_{n,k} z^n u^k$ ... bivariate GF of $(a_{n,k})_{n\geq 0, k\geq 0}$

$\Rightarrow [z^n u^k] A(z,u) = a_{n,k}$

- **Generating functions:**

$a_{n,k}$ ... number of lambda-terms of size n with k variables

$A(z,u) = \sum_{n,k\geq 0} a_{n,k} z^n u^k$ ... bivariate GF of $(a_{n,k})_{n\geq 0, k\geq 0}$

$\Rightarrow [z^n u^k] A(z,u) = a_{n,k}$

- **Singularity analysis:**

  ρ ... dominant singularity of $A(z)$ of type $-\alpha$ (*i.e.*, $(1-\frac{z}{\rho})^{-\alpha}$)

$$[z^n]A(z) \sim \rho^{-n} \frac{n^{\alpha-1}}{\Gamma(\alpha)} \qquad \text{as } n \longrightarrow \infty$$

- **Generating functions:**

$a_{n,k} \ldots$ number of lambda-terms of size n with k variables

$A(z,u) = \sum_{n,k\geq 0} a_{n,k} z^n u^k \ldots$ bivariate GF of $(a_{n,k})_{n\geq 0, k\geq 0}$

$\Rightarrow [z^n u^k] A(z,u) = a_{n,k}$

- **Singularity analysis:**

  $\rho \ldots$ dominant singularity of $A(z)$ of type $-\alpha$ (*i.e.*, $(1-\frac{z}{\rho})^{-\alpha}$)

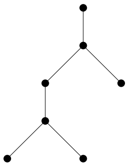
$$[z^n]A(z) \sim \rho^{-n} \frac{n^{\alpha-1}}{\Gamma(\alpha)} \qquad \text{as } n \longrightarrow \infty$$

- **Symbolic method:**

$$\mathcal{C} = \mathcal{A} \cup \mathcal{B} \qquad \Rightarrow C(z) = A(z) + B(z)$$
$$\mathcal{C} = \mathcal{A} \times \mathcal{B} \qquad \Rightarrow C(z) = A(z) \cdot B(z)$$

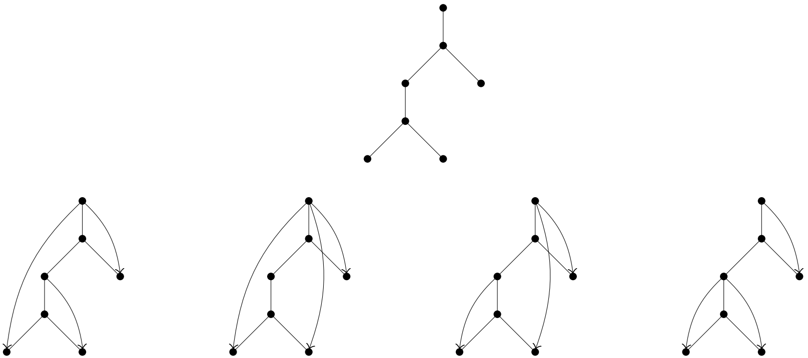
Reason for super exponential growth: possible variable bindings

$\rightarrow$ Comparison to Motzkin trees: $M_n = \sqrt{\frac{3}{4\pi n^3}}3^n + O(3^n n^{-2})$



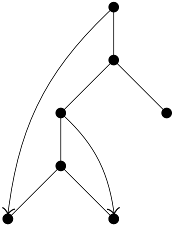
Reason for super exponential growth: possible variable bindings

$\rightarrow$ Comparison to Motzkin trees: $M_n = \sqrt{\frac{3}{4\pi n^3}} 3^n + O(3^n n^{-2})$

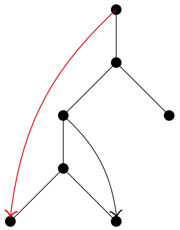


Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree

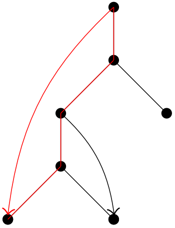
Unary length of a binding of a variable e by some abstraction $\lambda e =$
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree



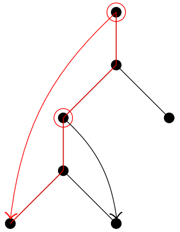
Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree



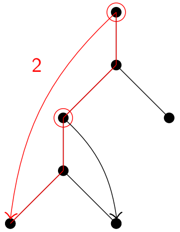
Unary length of a binding of a variable e by some abstraction $\lambda e =$
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree



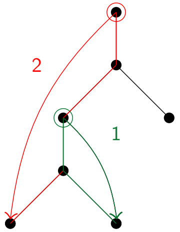
Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree



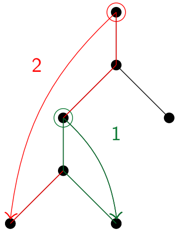
Unary length of a binding of a variable e by some abstraction $\lambda e =$
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree



Unary length of a binding of a variable e by some abstraction $\lambda e =$
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree

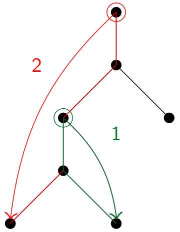


Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree

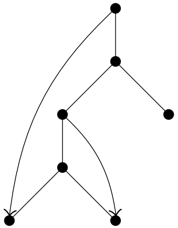


Unary height $h_u(e)$ of a vertex e =
number of unary nodes on the path connecting the root and e in the underlying Motzkin tree (the vertex itself does not count)

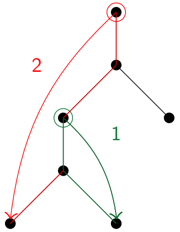
Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree



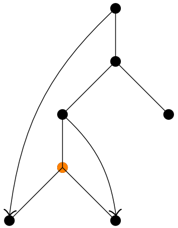
Unary height $h_u(e)$ of a vertex e =
number of unary nodes on the path connecting the root and e in the underlying Motzkin tree (the vertex itself does not count)



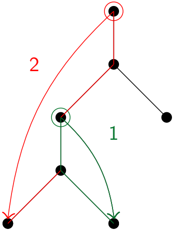
Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree



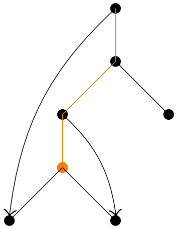
Unary height $h_u(e)$ of a vertex e =
number of unary nodes on the path connecting the root and e in the underlying Motzkin tree (the vertex itself does not count)



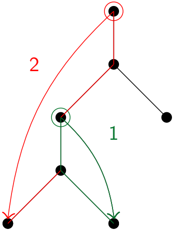
Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree



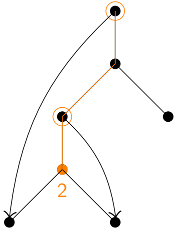
Unary height $h_u(e)$ of a vertex e =
number of unary nodes on the path connecting the root and e in the underlying Motzkin tree (the vertex itself does not count)



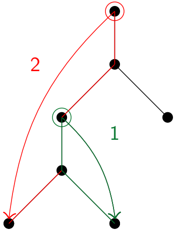
Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree



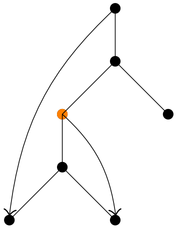
Unary height $h_u(e)$ of a vertex e =
number of unary nodes on the path connecting the root and e in the underlying Motzkin tree (the vertex itself does not count)



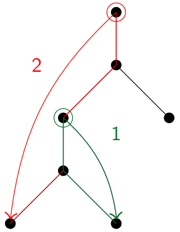
Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree



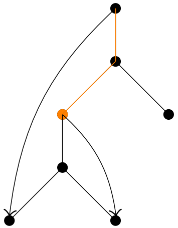
Unary height $h_u(e)$ of a vertex e =
number of unary nodes on the path connecting the root and e in the underlying Motzkin tree (the vertex itself does not count)



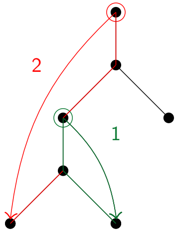
Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree



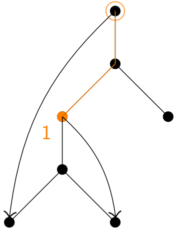
Unary height $h_u(e)$ of a vertex e =
number of unary nodes on the path connecting the root and e in the underlying Motzkin tree (the vertex itself does not count)



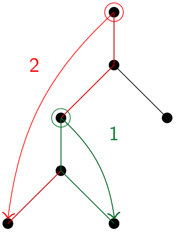
Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree



Unary height $h_u(e)$ of a vertex e =
number of unary nodes on the path connecting the root and e in the underlying Motzkin tree (the vertex itself does not count)

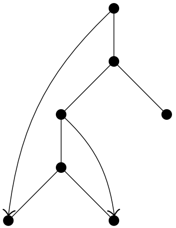


Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree

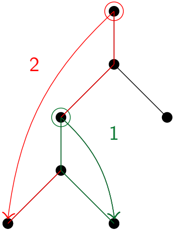


Unary height $h_u(e)$ of a vertex e =
number of unary nodes on the path connecting the root and e in the underlying Motzkin tree (the vertex itself does not count)

Unary height $h_u(t)$ of a lambda-term =
maximum number of unary nodes occurring in the branches of the underlying Motzkin tree

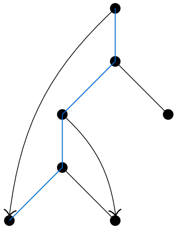


Unary length of a binding of a variable e by some abstraction λe = number of unary nodes on the path connecting λe and e in the underlying Motzkin tree

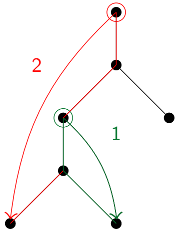


Unary height $h_u(e)$ of a vertex e = number of unary nodes on the path connecting the root and e in the underlying Motzkin tree (the vertex itself does not count)

Unary height $h_u(t)$ of a lambda-term = maximum number of unary nodes occurring in the branches of the underlying Motzkin tree

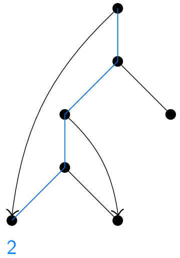


Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree

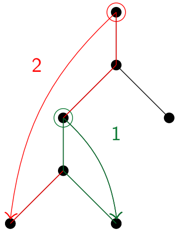


Unary height $h_u(e)$ of a vertex e =
number of unary nodes on the path connecting the root and e in the underlying Motzkin tree (the vertex itself does not count)

Unary height $h_u(t)$ of a lambda-term =
maximum number of unary nodes occurring in the branches of the underlying Motzkin tree

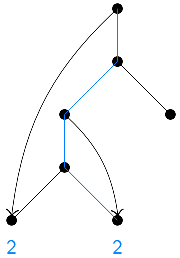


Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree

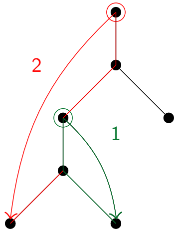


Unary height $h_u(e)$ of a vertex e =
number of unary nodes on the path connecting the root and e in the underlying Motzkin tree (the vertex itself does not count)

Unary height $h_u(t)$ of a lambda-term =
maximum number of unary nodes occurring in the branches of the underlying Motzkin tree

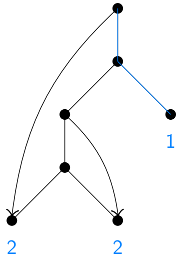


Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree

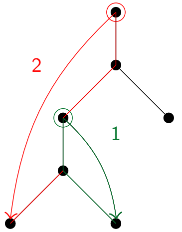


Unary height $h_u(e)$ of a vertex e =
number of unary nodes on the path connecting the root and e in the underlying Motzkin tree (the vertex itself does not count)

Unary height $h_u(t)$ of a lambda-term =
maximum number of unary nodes occurring in the branches of the underlying Motzkin tree



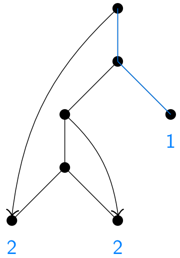
Unary length of a binding of a variable e by some abstraction λe =
number of unary nodes on the path connecting λe and e in the underlying Motzkin tree



Unary height $h_u(e)$ of a vertex e =
number of unary nodes on the path connecting the root and e in the underlying Motzkin tree (the vertex itself does not count)

Unary height $h_u(t)$ of a lambda-term =
maximum number of unary nodes occurring in the branches of the underlying Motzkin tree

$$h_u(t) = 2$$



Lambda-terms with bounded unary length of each binding

$\mathcal{G}_k \ldots$ class of closed lambda-terms where all bindings have unary length at most k

Lambda-terms with bounded unary length of each binding

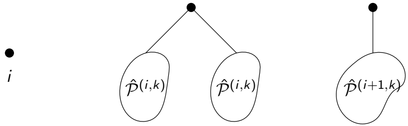
$\mathcal{G}_k \ldots$ class of closed lambda-terms where all bindings have unary length at most k

$\hat{\mathcal{P}}^{(i,k)} \ldots$ class of unary-binary trees such that every leaf e can be labeled in $\min\{h_u(e) + i, k\}$ ways

Lambda-terms with bounded unary length of each binding

$\mathcal{G}_k \ldots$ class of closed lambda-terms where all bindings have unary length at most k

$\hat{\mathcal{P}}^{(i,k)} \ldots$ class of unary-binary trees such that every leaf e can be labeled in $\min\{h_u(e)+i, k\}$ ways



$$\hat{\mathcal{P}}^{(k,k)} = k\mathcal{Z} + (\mathcal{A} \times \hat{\mathcal{P}}^{(k,k)} \times \hat{\mathcal{P}}^{(k,k)}) + (\mathcal{U} \times \hat{\mathcal{P}}^{(k,k)})$$

$$\hat{\mathcal{P}}^{(i,k)} = i\mathcal{Z} + (\mathcal{A} \times \hat{\mathcal{P}}^{(i,k)} \times \hat{\mathcal{P}}^{(i,k)}) + (\mathcal{U} \times \hat{\mathcal{P}}^{(i+1,k)}) \quad \text{for } i < k$$

$\mathcal{Z} \ldots$ class of atoms

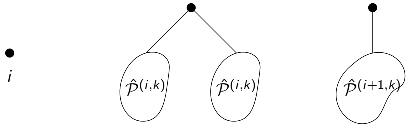
$\mathcal{A} \ldots$ class of application/binary nodes

$\mathcal{U} \ldots$ class of abstraction/unary nodes

Lambda-terms with bounded unary length of each binding

$\mathcal{G}_k \ldots$ class of closed lambda-terms where all bindings have unary length at most k

$\hat{\mathcal{P}}^{(i,k)} \ldots$ class of unary-binary trees such that every leaf e can be labeled in $\min\{h_u(e)+i, k\}$ ways



$$\hat{\mathcal{P}}^{(k,k)} = k\mathcal{Z} + (\mathcal{A} \times \hat{\mathcal{P}}^{(k,k)} \times \hat{\mathcal{P}}^{(k,k)}) + (\mathcal{U} \times \hat{\mathcal{P}}^{(k,k)})$$

$$\hat{\mathcal{P}}^{(i,k)} = i\mathcal{Z} + (\mathcal{A} \times \hat{\mathcal{P}}^{(i,k)} \times \hat{\mathcal{P}}^{(i,k)}) + (\mathcal{U} \times \hat{\mathcal{P}}^{(i+1,k)}) \quad \text{for } i < k$$

$\mathcal{Z} \ldots$ class of atoms

$\mathcal{A} \ldots$ class of application/binary nodes

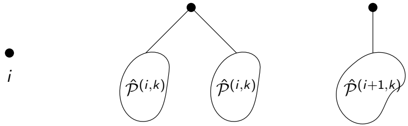
$\mathcal{U} \ldots$ class of abstraction/unary nodes

$$\boxed{\mathcal{G}_k \cong \hat{\mathcal{P}}^{(0,k)}}$$

Lambda-terms with bounded unary length of each binding

$\mathcal{G}_k \ldots$ class of closed lambda-terms where all bindings have unary length at most k

$\hat{\mathcal{P}}^{(i,k)} \ldots$ class of unary-binary trees such that every leaf e can be labeled in $\min\{h_u(e)+i,k\}$ ways



$$\hat{\mathcal{P}}^{(k,k)} = k\mathcal{Z} + (\mathcal{A} \times \hat{\mathcal{P}}^{(k,k)} \times \hat{\mathcal{P}}^{(k,k)}) + (\mathcal{U} \times \hat{\mathcal{P}}^{(k,k)})$$

$$\hat{\mathcal{P}}^{(i,k)} = i\mathcal{Z} + (\mathcal{A} \times \hat{\mathcal{P}}^{(i,k)} \times \hat{\mathcal{P}}^{(i,k)}) + (\mathcal{U} \times \hat{\mathcal{P}}^{(i+1,k)}) \quad \text{for } i < k$$

$$\updownarrow$$

$$\hat{P}^{(k,k)}(z) = kz + z\hat{P}^{(k,k)^2}(z) + z\hat{P}^{(k,k)}(z)$$

$$\hat{P}^{(i,k)}(z) = iz + z\hat{P}^{(i,k)^2}(z) + z\hat{P}^{(i+1,k)}(z)$$

$$\Rightarrow \hat{P}^{(i,k)}(z) = \frac{1-\mathbf{1}_{[i=k]}z-\sqrt{\hat{R}_{k-i+1,k}(z)}}{2z}$$

with

$$\hat{R}_{1,k}(z) = (1-z)^2 - 4kz^2,$$
$$\hat{R}_{2,k}(z) = 1 - 4(k-1)z^2 - 2z + 2z^2 + 2z\sqrt{\hat{R}_{1,k}(z)},$$

and for $3 \leq i \leq k+1$

$$\hat{R}_{i,k}(z) = 1 - 4(k-i+1)z^2 - 2z + 2z\sqrt{\hat{R}_{i-1,k}(z)}.$$

$$\boxed{\Rightarrow G_k(z) = \hat{P}^{(0,k)}(z) = \frac{1-\sqrt{\hat{R}_{k+1}(z)}}{2z}}$$

$$\Rightarrow \hat{P}^{(i,k)}(z) = \frac{1-\mathbf{1}_{[i=k]}z-\sqrt{\hat{R}_{k-i+1,k}(z)}}{2z}$$

with

$$\hat{R}_{1,k}(z) = (1-z)^2 - 4kz^2,$$
$$\hat{R}_{2,k}(z) = 1 - 4(k-1)z^2 - 2z + 2z^2 + 2z\sqrt{\hat{R}_{1,k}(z)},$$

and for $3 \leq i \leq k+1$

$$\hat{R}_{i,k}(z) = 1 - 4(k-i+1)z^2 - 2z + 2z\sqrt{\hat{R}_{i-1,k}(z)}.$$

$$G_k(z) = \frac{1-\sqrt{1-2z+2z\sqrt{\ldots 1-4(k-1)z^2-2z+2z^2+2z\sqrt{(1-z)^2-4kz^2}}}}{2z}$$

Theorem (Bodini, Gardy, Gittenberger, Gołębiewski, 2018)

- *The dominant singularity* $\hat{\rho}_k$ *of* $G_k(z)$ *comes from* $\hat{R}_1(z) = (1-z)^2 - 4kz^2$, *i.e.,* $\hat{\rho}_k = \frac{1}{1+2\sqrt{k}}$, *and is of type* $\frac{1}{2}$
- *For* $n \longrightarrow \infty$

$$[z^n]G_k(z) = g_k(1+2\sqrt{k})^n n^{-\frac{3}{2}} \left(1 + O\left(\frac{1}{n}\right)\right),$$

with $g_k = \sqrt{\frac{\sqrt{k}+2k}{4\pi \prod_{l=2}^{k+1} c_l}}$,

where $c_1 = 1$ *and* $c_j = 4(j-1) - 1 + 2\sqrt{c_{j-1}}$, *for* $2 \leq j \leq k+1$.

$$\boxed{G_k(z) = \hat{P}^{(0,k)}(z) = \frac{1 - \sqrt{\hat{R}_{k+1}(z)}}{2z}}$$

where

$$\hat{R}_{1,k}(z) = (1-z)^2 - 4kz^2,$$

$$\hat{R}_{2,k}(z) = 1 - 4(k-1)z^2 - 2z + 2z^2 + 2z\sqrt{\hat{R}_{1,k}(z)},$$

and for $3 \leq i \leq k+1$

$$\hat{R}_{i,k}(z) = 1 - 4(k-i+1)z^2 - 2z + 2z\sqrt{\hat{R}_{i-1,k}(z)}.$$

$$\boxed{G_k(z,u) = \hat{P}^{(0,k)}(z,u) = \frac{1-\sqrt{\hat{R}_{k+1}(z,u)}}{2z}}$$

where

$$\hat{R}_{1,k}(z,u) = (1-z)^2 - 4kuz^2,$$

$$\hat{R}_{2,k}(z,u) = 1 - 4(k-1)z^2u - 2z + 2z^2 + 2z\sqrt{\hat{R}_{1,k}(z,u)},$$

and for $3 \leq i \leq k+1$

$$\hat{R}_{i,k}(z,u) = 1 - 4(k-i+1)z^2u - 2z + 2z\sqrt{\hat{R}_{i-1,k}(z,u)}.$$

Theorem (Gittenberger, L., 2018)

- *The dominant singularity* $\hat{\rho}_k(u)$ *of* $G_k(z,u)$ *comes from* $\hat{R}_1(z,u) = (1-z)^2 - 4kuz^2$, *i.e.,* $\hat{\rho}_k(u) = \frac{1}{1+2\sqrt{ku}}$, *and is of type* $\frac{1}{2}$
- *For* $n \longrightarrow \infty$

$$[z^n]G_k(z,u) = g_k(u)(1+2\sqrt{ku})^n n^{-\frac{3}{2}} \left(1 + O\left(\frac{1}{n}\right)\right),$$

with $g_k(u) = \sqrt{\frac{\sqrt{ku}+2ku}{4\pi \prod_{l=2}^{k+1} c_l(u)}}$,

where $c_1(u) = 1$ *and* $c_j(u) = 4(j-1)u - 1 + 2\sqrt{c_{j-1}(u)}$, *for* $2 \le j \le k+1$.

Theorem (Quasi-Power Theorem, Hwang, 1998)

Let X_n be a sequence of random variables with the property that

$$\frac{[z^n]G(z,u)}{[z^n]G(z,1)} \sim A(u)B(u)^{\lambda_n},$$

with $\lambda_n \to \infty$.
Set $\mu = B'(1)$ and $\sigma^2 = B''(1) + B'(1) - B'(1)^2$. If $\sigma^2 \neq 0$, then

$$\frac{X_n - \mathbb{E}X_n}{\sqrt{\mathbb{V}X_n}} \to \mathcal{N}(0,1), \qquad \text{as } n \to \infty,$$

with

$$\mathbb{E}X_n \sim \mu\lambda_n \qquad \text{and} \qquad \mathbb{V}X_n \sim \sigma^2\lambda_n.$$

$$\frac{[z^n]G_k(z,u)}{[z^n]G_k(z,1)} = \left(\frac{1+2\sqrt{ku}}{1+2\sqrt{k}}\right)^n \sqrt{\frac{\sqrt{ku}+2ku}{2k+\sqrt{k}} \prod_{j=2}^{k+1} \frac{c_j(1)}{c_j(u)}} \left(1+O\left(\frac{1}{n}\right)\right)$$

$$\frac{[z^n]G_k(z,u)}{[z^n]G_k(z,1)} = \left(\frac{1+2\sqrt{ku}}{1+2\sqrt{k}}\right)^n \sqrt{\frac{\sqrt{ku}+2ku}{2k+\sqrt{k}} \prod_{j=2}^{k+1} \frac{c_j(1)}{c_j(u)}} \left(1+O\left(\frac{1}{n}\right)\right)$$

Theorem (Gittenberger, L., 2018)

Let X_n be the number of variables in closed lambda-terms where all bindings have unary length at most k. Then X_n is asymptotically normally distributed with asymptotic mean and variance given by

$$\mathbb{E}X_n \sim \frac{k}{\sqrt{k}+2k} n \quad \text{and} \quad \mathbb{V}X_n \sim \frac{k^2}{2\sqrt{k}(\sqrt{k}+2k)^2} n.$$

$$\frac{[z^n]G_k(z,u)}{[z^n]G_k(z,1)} = \left(\frac{1+2\sqrt{ku}}{1+2\sqrt{k}}\right)^n \sqrt{\frac{\sqrt{ku}+2ku}{2k+\sqrt{k}} \prod_{j=2}^{k+1} \frac{c_j(1)}{c_j(u)}} \left(1+O\left(\frac{1}{n}\right)\right)$$

Theorem (Gittenberger, L., 2018)

Let X_n be the number of variables in closed lambda-terms where all bindings have unary length at most k. Then X_n is asymptotically normally distributed with asymptotic mean and variance given by

$$\mathbb{E}X_n \sim \frac{k}{\sqrt{k}+2k} n \quad \text{and} \quad \mathbb{V}X_n \sim \frac{k^2}{2\sqrt{k}(\sqrt{k}+2k)^2} n.$$

For $k \longrightarrow \infty$: $\quad \mathbb{E}X_n \longrightarrow \frac{n}{2} \quad$ and $\quad \mathbb{V}X_n \longrightarrow 0.$

Lambda-terms with bounded unary height

$\mathcal{H}_k \ldots$ class of closed lambda-terms with unary height at most k

Lambda-terms with bounded unary height

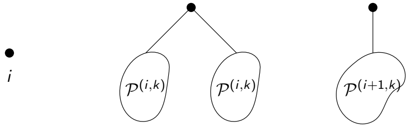
$\mathcal{H}_k \ldots$ class of closed lambda-terms with unary height at most k

$\mathcal{P}^{(i,k)} \ldots$ class of unary-binary trees such that every leaf e can be labeled in $\min\{h_u(e)+i, k\}$ ways and $h_u(e) \le k - i$.

Lambda-terms with bounded unary height

$\mathcal{H}_k \ldots$ class of closed lambda-terms with unary height at most k

$\mathcal{P}^{(i,k)} \ldots$ class of unary-binary trees such that every leaf e can be labeled in $\min\{h_u(e) + i, k\}$ ways and $h_u(e) \leq k - i$.



$$\mathcal{P}^{(k,k)} = k\mathcal{Z} + (\mathcal{A} \times \mathcal{P}^{(k,k)} \times \mathcal{P}^{(k,k)})$$
$$\mathcal{P}^{(i,k)} = i\mathcal{Z} + (\mathcal{A} \times \mathcal{P}^{(i,k)} \times \mathcal{P}^{(i,k)}) + (\mathcal{U} \times \mathcal{P}^{(i+1,k)}) \quad \text{for } i < k$$

$\mathcal{Z} \ldots$ class of atoms

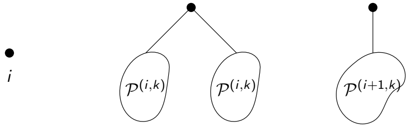
$\mathcal{A} \ldots$ class of application/binary nodes

$\mathcal{U} \ldots$ class of abstraction/unary nodes

Lambda-terms with bounded unary height

$\mathcal{H}_k \ldots$ class of closed lambda-terms with unary height at most k

$\mathcal{P}^{(i,k)} \ldots$ class of unary-binary trees such that every leaf e can be labeled in $\min\{h_u(e) + i, k\}$ ways and $h_u(e) \leq k - i$.



$$\mathcal{P}^{(k,k)} = k\mathcal{Z} + (\mathcal{A} \times \mathcal{P}^{(k,k)} \times \mathcal{P}^{(k,k)})$$

$$\mathcal{P}^{(i,k)} = i\mathcal{Z} + (\mathcal{A} \times \mathcal{P}^{(i,k)} \times \mathcal{P}^{(i,k)}) + (\mathcal{U} \times \mathcal{P}^{(i+1,k)}) \quad \text{for } i < k$$

$\mathcal{Z} \ldots$ class of atoms

$\mathcal{A} \ldots$ class of application/binary nodes

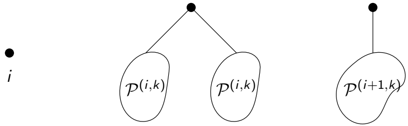
$\mathcal{U} \ldots$ class of abstraction/unary nodes

$$\boxed{\mathcal{H}_k \cong \mathcal{P}^{(0,k)}}$$

Lambda-terms with bounded unary height

$\mathcal{H}_k \ldots$ class of closed lambda-terms with unary height at most k

$\mathcal{P}^{(i,k)} \ldots$ class of unary-binary trees such that every leaf e can be labeled in $\min\{h_u(e)+i, k\}$ ways and $h_u(e) \le k-i$.



$$\mathcal{P}^{(k,k)} = k\mathcal{Z} + (\mathcal{A} \times \mathcal{P}^{(k,k)} \times \mathcal{P}^{(k,k)})$$

$$\mathcal{P}^{(i,k)} = i\mathcal{Z} + (\mathcal{A} \times \mathcal{P}^{(i,k)} \times \mathcal{P}^{(i,k)}) + (\mathcal{U} \times \mathcal{P}^{(i+1,k)}) \quad \text{for } i < k$$

$$\updownarrow$$

$$P^{(k,k)}(z) = kz + zP^{(k,k)^2}(z)$$

$$P^{(i,k)}(z) = iz + zP^{(i,k)^2}(z) + zP^{(i+1,k)}(z)$$

$$\Rightarrow P^{(i,k)}(z) = \frac{1-\sqrt{R_{k-i+1,k}(z)}}{2z}$$

with

$$R_{1,k}(z) = 1 - 4kz^2,$$

and for $2 \leq i \leq k+1$

$$R_{i,k}(z) = 1 - 4(k-i+1)z^2 - 2z + 2z\sqrt{R_{i-1,k}(z)}.$$

$$\boxed{\Rightarrow H_k(z) = P^{(0,k)}(z) = \frac{1-\sqrt{R_{k+1,k}(z,u)}}{2z}}$$

$$\Rightarrow P^{(i,k)}(z) = \frac{1-\sqrt{R_{k-i+1,k}(z)}}{2z}$$

with

$$R_{1,k}(z) = 1 - 4kz^2,$$

and for $2 \leq i \leq k+1$

$$R_{i,k}(z) = 1 - 4(k-i+1)z^2 - 2z + 2z\sqrt{R_{i-1,k}(z)}.$$

$$\boxed{H_k(z) = \frac{1 - \sqrt{1 - 2z + 2z\sqrt{\ldots 1 - 4(k-1)z^2 - 2z + 2z\sqrt{1-4kz^2}}}}{2z}}$$

Auxiliary sequences:

$$u_0 = 0, \quad u_{i+1} = u_i^2 + i + 1 \quad \text{for } i \geq 0,$$

and

$$N_i = u_i^2 - u_i + i, \quad \text{for all } i \geq 0.$$

j	N_j	u_j
1	1	1
2	8	3
3	135	12
4	21760	148
5	479982377	21909
⋮	⋮	⋮

Auxiliary sequences:

$$u_0 = 0, \quad u_{i+1} = u_i^2 + i + 1 \quad \text{for } i \geq 0,$$

and

$$N_i = u_i^2 - u_i + i, \quad \text{for all } i \geq 0.$$

j	N_j	u_j
1	1	1
2	8	3
3	135	12
4	21760	148
5	479982377	21909
$\vdots$	$\vdots$	$\vdots$

Theorem (Bodini, Gardy, Gittenberger, Gołębiewski, 2018)

Let ρ_k be the dominant singularity of $H_k(z)$.

(i) *If $N_j < k < N_{j+1}$, then ρ_k comes from R_{j+1} and*

$$[z^n]H_k(z) \sim h_k n^{-3/2} \rho_k^{-n}, \text{ as } n \to \infty.$$

(ii) *If $k = N_j$, then $\rho_k = \frac{1}{2u_j}$ and*

$$[z^n]H_k(z) \sim h_{N_j} n^{-5/4} \rho_k^{-n} = h_{N_j} n^{-5/4} (2u_j)^n, \text{ as } n \to \infty.$$

$$\boxed{H_k(z) = P^{(0,k)}(z) = \frac{1-\sqrt{R_{k+1}(z)}}{2z}}$$

where

$$R_{1,k}(z) = 1 - 4kz^2,$$

and for $2 \leq i \leq k+1$

$$R_{i,k}(z) = 1 - 4(k-i+1)z^2 - 2z + 2z\sqrt{R_{i-1,k}(z,u)}.$$

$$\boxed{H_k(z, u) = P^{(0,k)}(z, u) = \frac{1 - \sqrt{R_{k+1}(z, u)}}{2z}}$$

where

$$R_{1,k}(z, u) = 1 - 4kz^2 u,$$

and for $2 \leq i \leq k + 1$

$$R_{i,k}(z, u) = 1 - 4(k - i + 1)z^2 u - 2z + 2z\sqrt{R_{i-1,k}(z, u)}.$$

Total number of variables

Theorem (Gittenberger, L., 2018)

Let $B(u) = \frac{\rho_k(u)}{\rho_k(1)}$. *If* $B''(1) + B'(1) - B'(1)^2 \neq 0$, *then the total number* X_n *of variables in closed lambda-terms with unary height at most* k *is asymptotically normally distributed with*

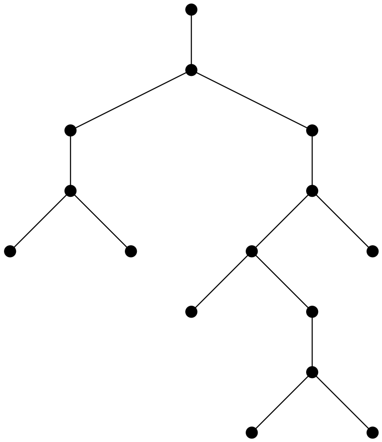
$$\mathbb{E}X_n \sim \mu \cdot n \qquad \text{and} \qquad \mathbb{V}X_n \sim \sigma^2 \cdot n,$$

where $\mu = B'(1)$ *and* $\sigma^2 = B''(1) + B'(1) - B'(1)^2$.

bound k	$j+1$	$B''(1)+B'(1)-B'(1)^2$	$B'(1)$
1	**2**	**0**	**0**
2	2	0.0385234386	0.4381229337
3	2	0.0210625856	0.4414407371
4	2	0.0167136805	0.4463973717
5	2	0.0148700270	0.4504258849
6	2	0.0138224393	0.4536185043
7	2	0.0131157948	0.4561987871
8	**3**	**0.048**	**0.4**
9	3	0.0582322465	0.4566104777
10	3	0.0470481360	0.4560418340
11	3	0.0396601986	0.4560810348
12	3	0.0345090124	0.4564489368
⋮	⋮	⋮	⋮
133	3	0.0077469541	0.4821900098
134	3	0.0077234960	0.4822482745
135	**4**	**0.0108490182**	**0.4782608696**

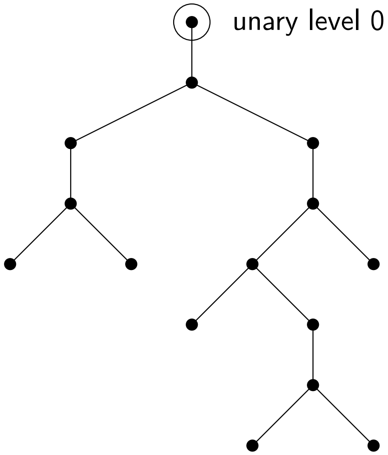
Number of variables in the separate unary levels

The *i*-**th unary level** contains all vertices with unary height *i*.



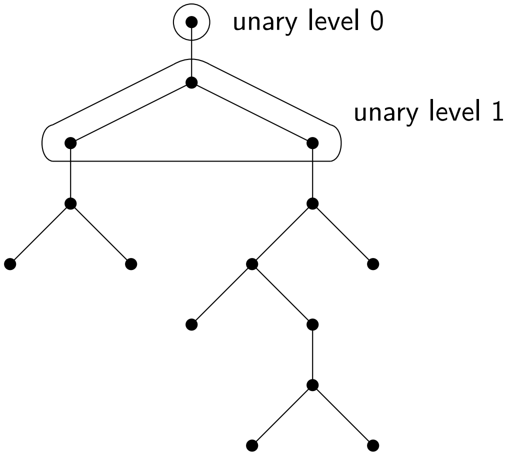
Number of variables in the separate unary levels

The *i*-**th unary level** contains all vertices with unary height *i*.



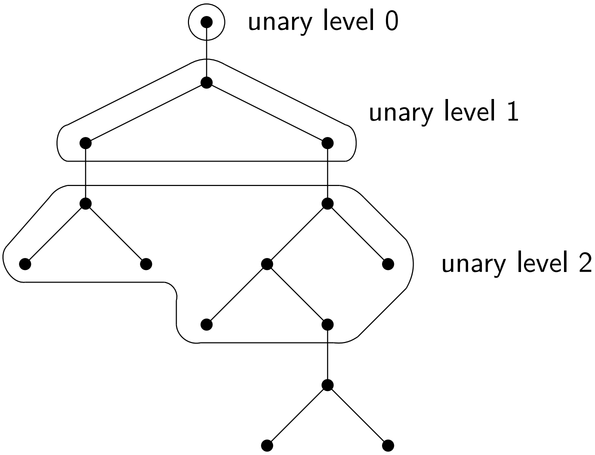
Number of variables in the separate unary levels

The *i*-**th unary level** contains all vertices with unary height i.



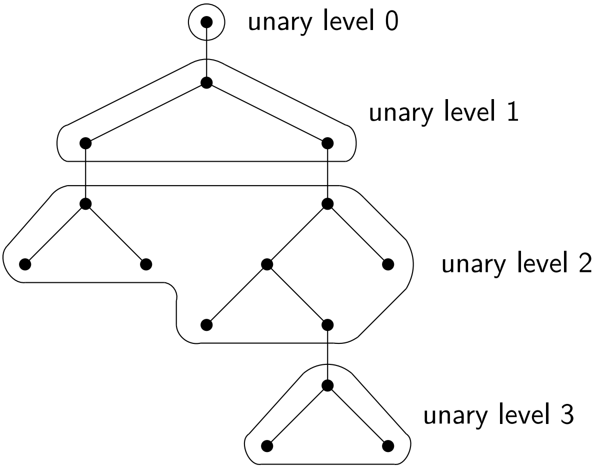
Number of variables in the separate unary levels

The *i*-**th unary level** contains all vertices with unary height i.



Number of variables in the separate unary levels

The *i*-**th unary level** contains all vertices with unary height i.



$C(z,u)\ldots$ GF of the class of binary trees where z marks the size and u the number of leaves $\to$ $C(z,u) = \frac{1-\sqrt{1-4uz^2}}{2z}$

$C(z,u)\ldots$ GF of the class of binary trees where z marks the size and u the number of leaves $\rightarrow C(z,u) = \frac{1-\sqrt{1-4uz^2}}{2z}$

Generating functions of lambda-terms with unary height at most k:

$H_k(z)\ldots$ z marks the size

$$H_k(z) = C(z, C(z, 1 + \ldots + C(z, (k-l) + \ldots + C(z, (k-1) + C(z,k))) \ldots) \ldots))$$

$C(z,u)\ldots$ GF of the class of binary trees where z marks the size and u the number of leaves $\to$ $C(z,u) = \frac{1-\sqrt{1-4uz^2}}{2z}$

Generating functions of lambda-terms with unary height at most k:

$H_k(z)\ldots$ z marks the size

$H_k(z,u)\ldots$ z marks the size, u marks the total number of leaves

$$H_k(z) = C(z, C(z, 1 + \ldots + C(z, (k-l) + \ldots + C(z, (k-1) + C(z,k)))\ldots)\ldots))$$

$$H_k(z,u) = C(z, C(z, u + \ldots + C(z, (k-l)\cdot u + \ldots + C(z, (k-1)\cdot u + C(z, k\cdot u)))\ldots)\ldots))$$

$C(z,u)\ldots$ GF of the class of binary trees where z marks the size and u the number of leaves $\rightarrow C(z,u) = \frac{1-\sqrt{1-4uz^2}}{2z}$

Generating functions of lambda-terms with unary height at most k:

$H_k(z)\ldots$ z marks the size

$H_k(z,u)\ldots$ z marks the size, u marks the total number of leaves

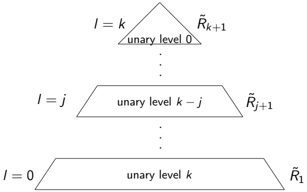
${}_{k-l}H_k(z,u)\ldots$ z marks the size, u marks the number of leaves on the $(k-l)$-th unary level $(0 \leq l \leq k)$

$$H_k(z) = C(z, C(z, 1 + \ldots + C(z, (k-l) + \ldots + C(z, (k-1) + C(z,k)))\ldots)\ldots))$$

$$H_k(z,u) = C(z, C(z, u + \ldots + C(z, (k-l)\cdot u + \ldots + C(z, (k-1)\cdot u + C(z, k\cdot u)))\ldots)\ldots))$$

$${}_{k-l}H_k(z,u) = C(z, C(z, 1 + \ldots + C(z, (k-l)\cdot u + \ldots + C(z, (k-1) + C(z,k)))\ldots)\ldots))$$

$$\boxed{{}_{k-l}H_k(z,u) = \frac{1-\sqrt{\tilde{R}_{k+1}(z,u)}}{2z}}$$



where

$$\tilde{R}_1(z,u) = 1 - 4z^2k,$$

and for $2 \le i \le k+1, i \neq l+1$

$$\tilde{R}_i(z,u) = 1 - 4z^2(k-i+1) - 2z + 2z\sqrt{\tilde{R}_{i-1}(z,u)},$$

and

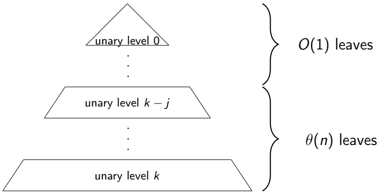
$$\tilde{R}_{l+1}(z,u) = 1 - 4z^2u(k-l) - 2z + 2z\sqrt{\tilde{R}_{l-1}(z,u)}.$$

Number of leaves on the separate unary levels

Let $k \in (N_j, N_{j+1})$

- $l > j : \mathbb{E}X_n \sim C_{k,l}$
- $l \leq j : \mathbb{E}X_n \sim \tilde{C}_{k,l} \cdot n$

with constants $C_{k,l}$ and $\tilde{C}_{k,l}$.

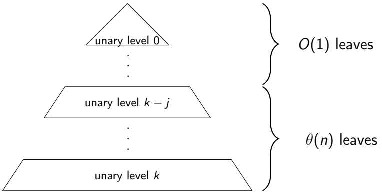


Number of leaves on the separate unary levels

Let $k \in (N_j, N_{j+1})$

- $l > j : \mathbb{E}X_n \sim C_{k,l}$
- $l \leq j : \mathbb{E}X_n \sim \tilde{C}_{k,l} \cdot n$

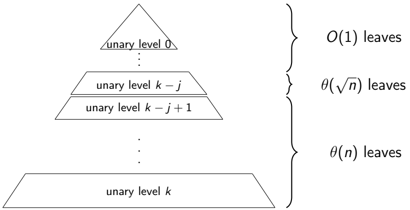
with constants $C_{k,l}$ and $\tilde{C}_{k,l}$.



Let $k = N_j$

- $l > j : \mathbb{E}X_n \sim D_{k,l}$
- $l = j : \mathbb{E}X_n \sim \hat{D}_{k,l} \cdot \sqrt{n}$
- $l < j : \mathbb{E}X_n \sim \tilde{D}_{k,l} \cdot n$

with constants $D_{k,l}$, $\hat{D}_{k,l}$ and $\tilde{D}_{k,l}$.

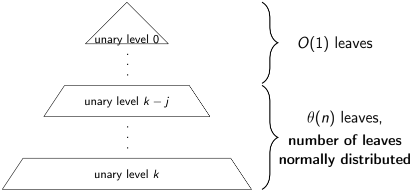


Number of leaves on the separate unary levels

Let $k \in (N_j, N_{j+1})$

- $l > j : \mathbb{E}X_n \sim C_{k,l}$
- $l \leq j : \mathbb{E}X_n \sim \tilde{C}_{k,l} \cdot n$

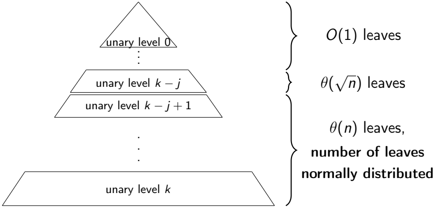
with constants $C_{k,l}$ and $\tilde{C}_{k,l}$.



Let $k = N_j$

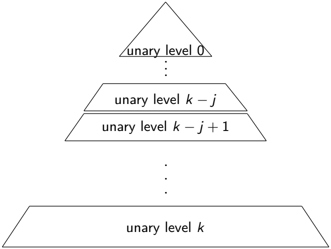
- $l > j : \mathbb{E}X_n \sim D_{k,l}$
- $l = j : \mathbb{E}X_n \sim \hat{D}_{k,l} \cdot \sqrt{n}$
- $l < j : \mathbb{E}X_n \sim \tilde{D}_{k,l} \cdot n$

with constants $D_{k,l}$, $\hat{D}_{k,l}$ and $\tilde{D}_{k,l}$.

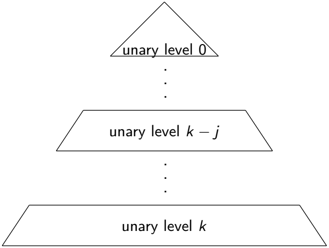


“Unary profile”

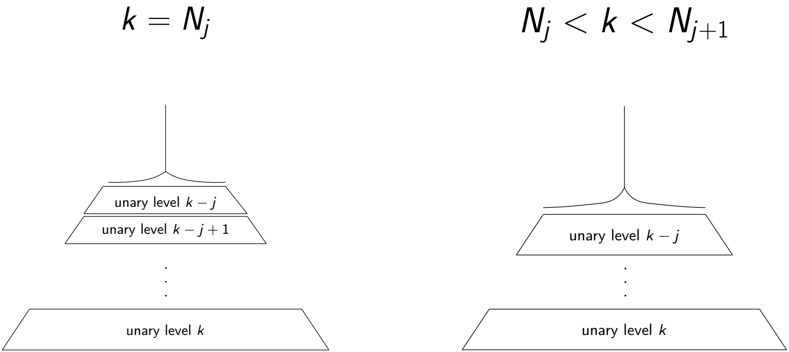
$$k = N_j$$



$$N_j < k < N_{j+1}$$



"Unary profile"



Thank you!