Enumeration and statistics of λ-terms

Danièle GARDY

PRiSM, Université Versailles St-Quentin en Yvelines
and
D.M.G., T.U. Wien

References

- *Asymptotics and random sampling for BCI and BCK lambda-terms*
  [Bodini G. Jacquot, Gascom 2010; journal version: TCS 2013]
- *Lambda-terms of bounded unary height*
  [Bodini G. Gittenberger, Analco 2011]
- *Enumeration of generalized BCI lambda-terms*
  [Bodini G. Gittenberger Jacquot, submitted Jan. 2013]
- *Work in progress*
  [Bodini G. Gittenberger 2014]

Syntactical approach to λ-terms

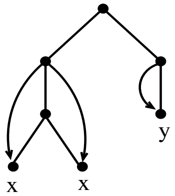
Extending quantitative logic

- Quantitative logic: study of random boolean expressions on several propositional logic systems, and of the induced probability distributions on boolean functions.
  Cf. Genitrini's talk
- Add quantifiers $\forall$ and $\exists$: extension of results?
- *Single* quantifier: expressions are equivalent to those of lambda-calculus
- Lambda-calculus has its own topics

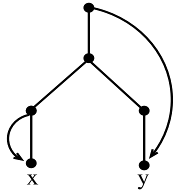
$$T ::= a \mid (T * T) \mid \lambda a.T$$

$(T * T)$: application $\lambda a.T$: abstraction

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

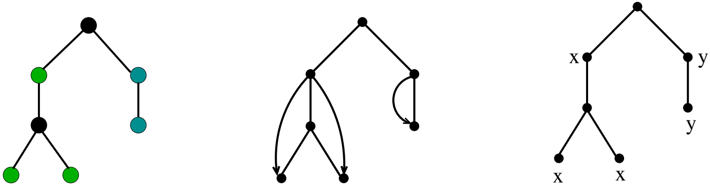


$\lambda y.(\lambda x.x * \lambda x.y)$



These λ-terms are closed (no free variable)

λ-terms as enriched Motzkin trees

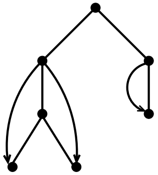


Labelling rules:

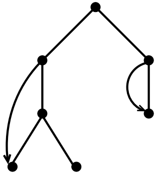
- Binary nodes are unlabelled
- Unary nodes get distinct labels (colors)
- Leaves get the label (color) of one of their unary ancestors

Free and bound variables in leaves

- Here all variables are bound



- Some variables may be free



- Recursive definition for λ-terms?
  - $\mathcal{L}$: class of λ-terms with free variables
  - $\mathcal{N}$ atomic class of binary node
  - $\mathcal{U}$ atomic class of unary node
  - $\mathcal{F}$ atomic class of free leaf
  - $\mathcal{B}$ atomic class of bound leaf

$$\mathcal{L} = \mathcal{F} + \left(\mathcal{N} \times \mathcal{L}^2\right) + (\mathcal{U} \times subs(\mathcal{F} \to \mathcal{F} + \mathcal{B}, \mathcal{L}))$$

- Generating function

$$L(z, f) = fz + z\,L(z, f)^2 + z\,L(z, f + 1).$$

with $z \leftrightarrow$ size of the λ-term and $f \leftrightarrow$ free leaves
(size = total number of nodes)

- Generating function enumerating closed λ-terms (without free variables): $L(z,0)$
- Generating function enumerating all λ-terms: $L(z,1) = \frac{1}{z}L(z,0) - L(z,0)^2$
- $L(z,0) = \frac{1}{2z}\left(1 - \sqrt{\Lambda(z)}\right)$ with $\Lambda(z)$ equal to

$$1 - 2z + 2z\sqrt{1 - 2z - 4z^2 + 2z\sqrt{\dots.\sqrt{1 - 2z - 4nz^2 + 2z\sqrt{\dots}}}}$$

- $L(z,0)$ has null radius of convergence $\Rightarrow$ standard tools of analytic combinatorics fail

What can we do?

- Try to find a way to deal with null radius of convergence?
- *Ad hoc* methods?

$$\left(\frac{(4-\epsilon)n}{\log n}\right)^{n(1-1/\log n)} \leq L_n \leq \left(\frac{(12+\epsilon)n}{\log n}\right)^{n(1-1/3\log n)}$$

[David et al. 10; here leaves have size 0]

- Consider sub-classes of terms?
  - Restrict the *total* number of abstractions
  - Restrict the number of abstractions *in a path from the root towards a leaf:* bounded unary height
  - Restrict the number of pointers from an abstraction to a leaf

λ-terms with bounded number of unary nodes

q unary nodes

$$\mathcal{S}_q = \left(\mathcal{U} \times subs(\mathcal{F} \to \mathcal{F} + \mathcal{B}, \mathcal{S}_{q-1})\right) + \sum_{\ell=0}^{q} (\mathcal{A}, \mathcal{S}_\ell, \mathcal{S}_{q-\ell})$$

Generating function

$$S_q(z, f) = zS_{q-1}(z, f + 1) + z \sum_{\ell=0}^{q} S_\ell(z, f)\ S_{q-\ell}(z, f).$$

G.F. for closed terms $S_q(z, 0)$?

$$S_1(z, 0) = \frac{1}{2} - \frac{\sqrt{1 - 4z^2}}{2};$$

$$S_2(z, 0) = \frac{z}{2}(1 - 2z^2) + \frac{2z^3}{\sqrt{1 - 4z^2}} - \frac{z\sqrt{1 - 8z^2}}{2\sqrt{1 - 4z^2}};$$

(no terms of size $n = q \ mod\ 2$)

q unary nodes

$$S_q(z,f) = -\frac{z^{q-1}\sigma_q(f)}{2\prod_{\ell=0}^{q-1}\sigma_\ell(f)} + R_q(z,\sigma_0(f),\ldots,\sigma_{q-1}(f))$$

where

- $\sigma_q(f) = \sqrt{1-4(f+q)z^2}$
- R_q rational, denominator $\prod_{0\leq\ell<q}\sigma_\ell(f)^{\alpha_{\ell,q}}$
- $\alpha_{\ell,q} > 0$, either integer or $\frac{1}{2}+$ an integer

q unary nodes

$$S_q(z,f) = -\frac{z^{q-1}\sigma_q(f)}{2\prod_{\ell=0}^{q-1}\sigma_\ell(f)} + R_q(z,\sigma_0(f),\dots,\sigma_{q-1}(f))$$

where

- $\sigma_q(f) = \sqrt{1-4(f+q)z^2}$
- R_q rational, denominator $\prod_{0\le\ell<q}\sigma_\ell(f)^{\alpha_{\ell,q}}$
- $\alpha_{\ell,q} > 0$, either integer or $\frac{1}{2}+$ an integer

$$\Rightarrow S_q(z,0) = -\frac{z^{q-1}\sqrt{1-4qz^2}}{2\prod_{\ell=0}^{q-1}\sqrt{1-4\ell z^2}} + R_q(z,1,\sqrt{1-4z^2},\dots,\sqrt{1-4(q-1)z^2})$$

q unary nodes

$$S_q(z,0) = -\frac{z^{q-1}\sqrt{1-4qz^2}}{2\prod_{\ell=0}^{q-1}\sqrt{1-4\ell z^2}} + R_q(z,1,\sqrt{1-4z^2},\ldots,\sqrt{1-4(q-1)z^2})$$

Dominant singularities at $\pm\frac{1}{2\sqrt{q}}$ of square-root type

$$\Rightarrow [z^n]S_q(z,0) \sim \frac{q^{\frac{q}{2}}}{\sqrt{q!}\sqrt{2\pi n^3}}(4q)^{\frac{n+1-q}{2}}$$

λ-terms of bounded unary height

The classes $\mathcal{P}^{(i,k)}$

k: maximal number of abstractions on a path from the root to a leaf

- $\mathcal{P}^{(0,k)}$: λ-terms with bound variables and unary height $\leq k$
- $\mathcal{P}^{(1,k)}$: λ-terms with bound variables, 1 kind of free variables, and unary height $\leq k - 1$
- ...
- $\mathcal{P}^{(i,k)}$: λ-terms with bound variables, i kinds of free variables, and unary height $\leq k - i$
- ...
- $\mathcal{P}^{(k,k)}$: λ-terms with bound variables, k kinds of free variables, and no unary node

The classes $\mathcal{P}^{(i,k)}$

- $\underline{i = k}$

$$\mathcal{P}^{(k,k)} = k\mathcal{Z} + \mathcal{Z}{\mathcal{P}^{(k,k)}}^2$$

Generating function:

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

- $\underline{i < k}$

$$\mathcal{P}^{(i,k)} = i\mathcal{Z} + \mathcal{Z}{\mathcal{P}^{(i,k)}}^2 + \mathcal{Z}\mathcal{P}^{(i+1,k)}$$

Generating function:

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

Solve in $P^{(i,k)}$ and take $H_{\leq k}(z) = P^{(0,k)}(z)$:

$$H_{\leq k} = \frac{1 - \sqrt{1 - 2z + 2z\sqrt{1 - 2z - 4z^2 + 2z\sqrt{\ldots\sqrt{1 - 4kz^2}}}}}{2z}$$

We can start the asymptotic study of its coefficients!

Solve in $P^{(i,k)}$ and take $H_{\leq k}(z) = P^{(0,k)}(z)$:

$$H_{\leq k} = \frac{1 - \sqrt{1 - 2z + 2z\sqrt{1 - 2z - 4z^2 + 2z\sqrt{\ldots\sqrt{1 - 4kz^2}}}}}{2z}$$

We can start the asymptotic study of its coefficients!

- $H_{\leq k}$ is algebraic and written with $k + 1$ iterated radicands
- Its singularities are the values that cancel its radicands
- Which radicant has smallest root?

- $k = 1$

$$H_{\leq 1}(z) = \frac{1 - \sqrt{1 - 2z + 2z\sqrt{1-4z^2}}}{2z}$$

Dominant singularity: $\frac{1}{2}$ (cancels both radicands)

$$[z^n]H_{\leq 1}(z) \sim \frac{1}{4} \frac{2^{\frac{1}{4}} 2^n n^{-\frac{5}{4}}}{\Gamma(\frac{3}{4})}$$

- $k = 2$

$$H_{\leq 2}(z) = \frac{1 - \sqrt{1 - 2z + 2z\sqrt{1 - 2z - 4z^2 + 2z\sqrt{1-8z^2}}}}{2z}$$

Dominant singularity: $\rho = 0.3437999303$ (cancels the second innermost radicand)

$$[z^n]H_{\leq 2}(z) \sim \frac{C}{\Gamma(\frac{1}{2})} n^{-\frac{3}{2}} \rho^{-n}$$

Where is the dominant singularity when k grows?

Function	Radicand	Singularity
$H_{\leq 1}$	{1,2}	0.5
$H_{\leq 2}$	2	0.3438
$H_{\leq 3}$	2	0.2760
...	...	...
$H_{\leq 8}$	{2,3}	0.1667
$H_{\leq 9}$	3	0.1571
...	...	...
$SH_{\leq 134}$	3	0.0418
$H_{\leq 135}$	{3,4}	0.0417
$H_{\leq 136}$	4	0.0415
...	...	...

Sometimes, the same value cancels *two* consecutive radicands.

Values of k which give two dominant radicands?

- Define $(u_k)_{k\geq 0}$: $u_0 = 0$ and $u_k = u_{k-1}^2 + k$ for $k > 0$
- First values: $u_1 = 1$, $u_2 = 3$, $u_3 = 12$, $u_4 = 148$, $u_5 = 21909$, ...
- The sequence $(u_k)_{k\geq 0}$ is doubly exponential
- $\lim_{k\to\infty} u_k^{1/2^k} \simeq \chi = 1.36660956...$
- Define $N_k = u_k^2 - u_k + k = u_k^2 - u_{k-1}^2$.
  $N_1 = 1$, $N_2 = 8$, $N_3 = 135$, $N_4 = 21760$, $N_5 = 479982377$, ...

Theorem

Two asymptotic behaviours according to the value of k

- *For unary height N_i, the radicands with ranks i and $(i+1)$ both cancel for the same value; both are dominant; the dominant singularity is algebraic of type $1/4$, and $[z^n]H_{\leq N_i} \sim C_i n^{-5/4} \rho_i^n$, with $\rho_i = 1/2u_i$.*
- *If $k \in]N_i, N_{i+1}[$, the dominant radicand of $H_{\leq k}(z)$ is the i-th radicand; the dominant singularity is algebraic of type $1/2$, and $[z^n]H_{\leq k} \sim C_k n^{-3/2} \rho_k^n$.*

(Radicands are ranked from the innermost to the outermost)

Observations

- The constants C_k become small very quickly.
  Variation of C_k as a function of k?
  - $[z^n]H_{\leq 1}(z) \sim 0.2426128012 \cdot \left(\frac{1}{n}\right)^{5/4} \cdot 2^n$
  - $[z^n]H_{\leq 8}(z) \sim 9.318885373 \cdot 10^{-5} \left(\frac{1}{n}\right)^{5/4} 6^n$
  - $[z^n]H_{\leq 135}(z) \sim 7.116999389 \cdot 10^{-158} \left(\frac{1}{n}\right)^{5/4} 24^n$
  - Doubly-exponential decay?
- We cannot hope to observe the asymptotic behaviour of $[z^n]H_{\leq k}$ from computations for “reasonable” n
- Yet we can randomly generate lambda-terms of bounded height and observe the behaviour of parameters...

The constant C_k for $k \in \{N_i\}$

$$[z^n]\mathcal{S}^{(N_k)} \sim \frac{D_k}{\Gamma(3/4)\, A_k} n^{-5/4} (2u_k)^n$$

where

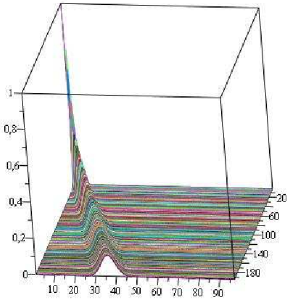
$$\begin{aligned}
D_k &= \left(-\frac{u_k}{2} \frac{d}{dz} R_{k,N_k}(\rho_k) \right)^{1/4} \\
R_{1,p}(z) &= 1 - 4pz^2 \\
R_{k,p}(z) &= 1 - 4(p-k+1)z^2 - 2z + 2z\sqrt{R_{k-1,p}(z)} \\
A_k &= \prod_{i=k}^{N_k - 1} \sqrt{i + \sqrt{i - 1 + \sqrt{i - 2 + \sqrt{\ldots + \sqrt{1}}}}}
\end{aligned}$$

Asymptotics for $k \to +\infty$?

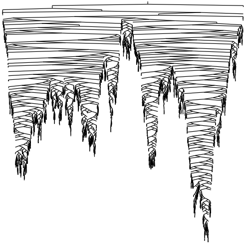
$$D_k \sim \gamma\sqrt{u_k} \sim \gamma u_{k-1} \qquad \textit{with} \qquad \gamma = 1.2952778$$

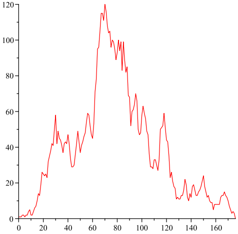
$$A_k \sim \frac{\varphi(N_k)}{\varphi(k)} \qquad \textit{with} \qquad \varphi(k) = \frac{e^{\sqrt{k}}}{\sqrt{k}} \cdot \left(\frac{2k}{e}\right)^k$$

Number of λ-terms: $n \in [1, ..., 198]$; unary height $k \in [1, ..., 98]$



A random λ-term of unary height ≤ 8 and its profile





λ-terms of bounded arity

λ-terms with fixed arity

Two classes of closed λ-terms:

- *BCI*(p): each abstraction binds *exactly* p variables (*linear* terms)
- *BCK*(p): each abstraction binds *at most* p variables (*affine* terms)

Consider first $p = 1$, then generalize...

$BCI(1)$ and $BCK(1)$

Class of λ-terms when each abstraction binds exactly one variable: $BCI(1)$

- Size is always $3n+2$
- Bijection with triangular pointed diagrams enumerated according to the number of edges (Vidal)
- Equate number of terms with coefficients of $z^3 \frac{d}{dz} \ln \left(e^{z^3/3} \odot e^{z^2/2} \right)$
- Get asymptotic equivalent $BCI(1)_{3n+2} \sim C\sqrt{n}\left(\frac{6n}{e}\right)^n$

Extension of approach for $BCI(1)$ gives

$$BCK(1)_n \sim \frac{C_1}{n^{1/6}} \left(\frac{2n}{e} \right)^{n/3} e^{\frac{(2n)^{2/3}}{2} - \frac{(2n)^{1/3}}{6}}$$

How can we generalize this approach to BCI(p)? to BCK(p)? to unrestricted λ-terms?

How can we generalize this approach to BCI(p)? to BCK(p)? to unrestricted λ-terms?

- Generalize the differential equation on the bivariate generating function for $BCI(1)$ terms with free leaves?

$$T(z,f) = zf + zT^2(z,f) + z\frac{\partial T}{\partial f}(z,f)$$

How can we generalize this approach to BCI(p)? to BCK(p)? to unrestricted λ-terms?

- Generalize the differential equation on the bivariate generating function for $BCI(1)$ terms with free leaves?

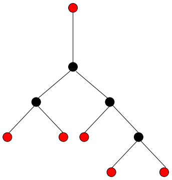
$$T(z,f) = zf + zT^2(z,f) + z\frac{\partial T}{\partial f}(z,f)$$

- Go back to the recursive definition of (unrestricted) λ-terms and adapt?

$$\mathcal{L} = \mathcal{F} + \left(\mathcal{N} \times \mathcal{L}^2\right) + (\mathcal{U} \times subs(\mathcal{F} \to \mathcal{F} + \mathcal{B}, \mathcal{L}))$$

Recursive construction of a *BCI*(p) term

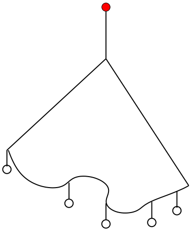
- A *BCI*(p) term with j abstraction nodes has size $(2p+1)j-1$
- Smallest terms: $j = 1$. There are C_{p-1} such terms



- All other terms are obtained either by taking two terms as left and right children of a binary root, or by taking a term and adding an abstraction node at root

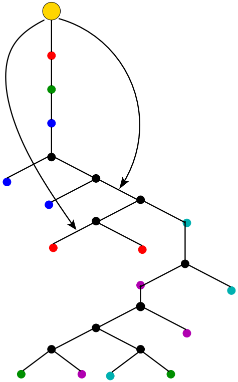
Adding an abstraction as root of a *BCI*(*p*) term

- In the term whose root is the new abstraction node, all other abstraction nodes already have *p* pointer to leaves
- The term below the root must have *p* free leaves...



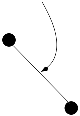
- ... but a *BCI*(*p*) term is closed!

How do we get new, free leaves?



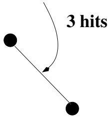
How do we get new, free leaves?

- If an edge is hit once...

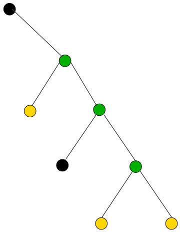


... we create a new leaf; there are 2 choices

- If an edge is hit *i* times....



... we create *i* new leaves and *i* new binary nodes. There are $\binom{2i}{i}$ ways to do it (build a sequence of binary trees with a total of *i* leaves; each new tree goes either right or left)



The differential operator Δ_p

- p hits
- some edges can be hit repeatedly
- l different edges are hit

$$\alpha_{l,p} = \sum_{\sum_i s_i = l; \sum_i i\, s_i = p} \binom{l}{s_1! \ldots s_p!} \prod_{m=1}^{p} \binom{2m}{m}^{s_m}$$

$$\Delta_p = \sum_{1 \leq l \leq p} \frac{\alpha_{l,p}}{l!} z^{l+2p+1} D^l$$

Generating function for $BCI(p)$

A $BCI(p)$ term either

1. is a Catalan tree with p leaves, topped by an abstraction node,
2. or has a binary node as root with two $BCI(p)$ children ,
3. or has abstraction node as root, pointing to p *new* free leaves

Univariate generating function for $BCI(p)$ is solution of a differential equation:

$$Y(z) = C_{p-1} z^{2p} + zY(z)^2 + \Delta_p Y(z)$$

What about $BCK(p)$?

Generating function $F_p(z) = Y\left(\frac{z}{1-z}\right)$, with Y satisfying

$$Y(z) = \sum_{l=1}^{p} C_{l-1} z^{2l} + zY(z)^2 + \left(\sum_{l=1}^{p} \Delta_p\right) Y(z)$$

Alternative form:

$$M(z,u) = \left(1 - z - \sqrt{(1-z)^2 - 4uz}\right)/(2z);$$

$$F_p(z) = z[u^p]\frac{M(z,u)}{1-u} + zF_p(z)^2 + z[u^p]\frac{1}{1-u} F_p\left(\frac{z}{1-2zM(z,u)}\right)$$

And for unrestricted λ-terms?

Generating function for closed terms is $\Lambda(z) = \tilde{\Lambda}\left(\frac{z}{1-z}\right)$ with

$$\tilde{\Lambda}(z) = C(z) + z\tilde{\Lambda}(z) + z\tilde{\Lambda}\left(\frac{z}{1-2C(z)}\right) - z\tilde{\Lambda}(z)$$

and $C(z) = (1 - \sqrt{1-4z^2})/2$

Alternative form:

$$\Lambda(z) = zM(z,1) + z\Lambda(z)^2 + z\Lambda\left(\frac{z}{1-2zM(z,1)}\right)$$

Solving the differential equation for $BCI(p)$?

$$Y = C_{p-1} z^{2p} + zY^2 + \Delta_p Y$$

We cannot solve explicitly this differential equation, nor find asymptotics by singularity analysis (radius of convergence is null again)...

Solving the differential equation for $BCI(p)$?

$$Y = C_{p-1} z^{2p} + zY^2 + \Delta_p Y$$

We cannot solve explicitly this differential equation, nor find asymptotics by singularity analysis (radius of convergence is null again)...
... but we can do asymptotics for an *approximate* equation

$$Y = C_{p-1} z^{2p} + 2C_{p-1} zY + \Delta_p Y$$

with same asymptotic behaviour!

Asymptotic enumeration of $BCI(p)$

Theorem

For $p \geq 2$, the number of λ-terms of $BCI(p)$ of size $(2p+1)n-1$ is asymptotically

$$a_p \, B_p \, \beta_p^{n-1} \, n^{\frac{p(p-2)}{2p+1}} \, n^{np}$$

with $a_p = 1 + \mathcal{O}(1/(pe^p))$, $\beta_p = \frac{(4p+2)^p}{p!}$ *and*

$$\begin{aligned} B_p &= \frac{C_{p-1}}{\prod_{1\leq j\leq p} \Gamma\left(1 + \frac{2(p-j)-1}{2p+1}\right)} \\ &\sim C_{p-1} \, (1.0844375...)^{(2p+1)/2} \, (1 + \mathcal{O}(1/p)). \end{aligned}$$

What about larger classes of terms?

Number λ_n of unrestricted λ-terms

Theorem

For all ε and for large n

$$c_1 \left(\frac{4n}{e\log n}\right)^{n/2} \frac{\sqrt{\log n}}{n} \leq \lambda_n \leq c_2 \left(\frac{9(1+\varepsilon)n}{e\log n}\right)^{n/2} \frac{(\log n)^{n/(2\log n)}}{n^{3/2}}$$

[Proof: by counting a subclass, and a larger set]

Recurrence relations for λ_n

$$\lambda_n = M_{n-1} + \sum_{\ell+q=n-1} \lambda_\ell \lambda_q + \sum_{1\leq\ell\leq n-1} \delta_{n,\ell}\lambda_\ell$$

D-finite recurrences for $\delta_{n,\ell}$:

$$\begin{aligned}
&(n-\ell)(n+1-\ell)(n-2\ell-2)\,\delta_{n+2,\ell}\\
&-(n-\ell)\left(2n^2-6n\ell-5n+2\ell^2+3\ell+1\right)\delta_{n+1,\ell}\\
&-(n-1)\left(3n^2-2n\ell+n-\ell^2-9\ell-8\right)\delta_{n,\ell}\\
&+20\,(n-1)\,\ell\,(\ell+1)\,\delta_{n,\ell+2}\\
&+2\,(n-1)(5n-9\ell-12)\,\ell\delta_{n,\ell+1}=0
\end{aligned}$$

and

$$\begin{aligned}
(n-\ell)(\ell-n-1)\,\delta_{n+2,\ell}+(n-\ell)(2n-\ell)\,\delta_{n+1,\ell}-\ell\,(n-1)\,\delta_{n+1,\ell+1}\\
-4\ell\,(n-1)\,\delta_{n,\ell+1}\,(n-1)(3n-2\ell+1)\,\delta_{n,\ell}=0.
\end{aligned}$$

Some properties of λ-terms?

λ-terms asymptotics differ widely from that of trees

1. Motzkin trees
   - Number of unary nodes = q: one radical, $C_q\, 4^n n^{q-\frac{3}{2}}$
   - Shared unary height of leaves = k: iterated radicals; *innermost* radical dominates; $C_k\, 2^n\, n^{-1-\frac{1}{2^{k+1}}}$
   - Bounded unary height = k: iterated radicals, *outermost* radical dominates; $C_k\, \rho_k^n\, n^{-\frac{3}{2}}$
2. λ-terms
   - Number of unary nodes = q: product of radicals; $C_q\, (4q)^{\frac{n+1-q}{2}}\, n^{-\frac{3}{2}}$
   - Bounded unary height = k: iterated radicals; dominant radical *fluctuates*
     - Standard case: $C_k n^{-\frac{3}{2}} \rho_k^n$
     - Special values: *two* dominant radicals; $C_k n^{-\frac{5}{4}} \rho_k^n$

Summary: number of λ-terms for several classes

- q abstractions: $\frac{q^{\frac{q}{2}}}{\sqrt{q!}\,\sqrt{2\,\pi\,n^3}}\,(4q)^{\frac{n+1-q}{2}}$
- Bounded unary height h: (usually) $C\rho_h^{-n}\,n^{-3/2}$
- $BCI(1)_{3n+2} \sim C\sqrt{n}\left(\frac{6n}{e}\right)^n$
- $BCI(p)_{(2p+1)n-1} \sim a_p\,B_p\,\beta_p^{n-1}\,n^{\frac{p(p-2)}{2p+1}}\,n^{np}$
- Unrestricted terms:

$$c_1\left(\frac{4n}{e\log n}\right)^{n/2}\frac{\sqrt{\log n}}{n} \le \lambda_n \le c_2\left(\frac{9(1+\varepsilon)n}{e\log n}\right)^{n/2}\frac{(\log n)^{\frac{n}{2\log n}}}{n^{3/2}}$$

Number of unary nodes in terms of bounded height k ?

Two different limiting distributions

- *k* is special: characteristic function is a variation on Bessel functions
- standard case: (probably) gaussian

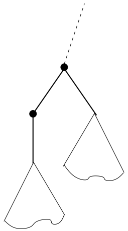
(work in progress)

λ-terms and normal form

A term is in normal form

$\Updownarrow$

there is no pattern $\mathcal{A} \times (\mathcal{U}, \mathcal{T}) \times \mathcal{T}$



Forbidden pattern

Number of terms in normal form?

Normal form, bounded number of unary nodes

- Asymptotic number of closed, normal-form λ-terms with exactly q unary nodes and size n, $n \neq q \mod 2$

$$\frac{1}{2^q \sqrt{2\pi n^3}} \prod_{\ell=1}^{q} \frac{\sqrt{q}+\sqrt{\ell}}{\sqrt{\ell}} (4q)^{\frac{n+1-q}{2}}$$

- Asymptotic probability of closed, normal-form term with exactly q unary nodes and size n ($n \to +\infty$)

$$\pi_q = 2^{-q} \prod_{\ell=1}^{q} \left(1 + \sqrt{\frac{\ell}{q}}\right)$$

Normal form, bounded number of unary nodes

Asymptotic probability of closed normal-form term with exactly q unary nodes and size n for large q

$$\pi_q = \sqrt{2}\left(\frac{\sqrt{e}}{2}\right)^q (1 + o(1)) = \sqrt{2}\ \ 0.82436^q\ (1 + o(1))\,.$$

q	5	10	50	100	1000
Exact	0.496	0.193	$8.79\ 10^{-5}$	$5.67\ 10^{-9}$	$1.84\ 10^{-84}$
Large q	0.538	0.205	$9.04\ 10^{-5}$	$5.78\ 10^{-9}$	$1.85\ 10^{-84}$

Normal form, bounded number of unary nodes

Asymptotic probability of closed normal-form term with exactly q unary nodes and size n for large q

$$\pi_q = \sqrt{2}\left(\frac{\sqrt{e}}{2}\right)^q (1 + o(1)) = \sqrt{2}\ \ 0.82436^q\ (1 + o(1))\,.$$

q	5	10	50	100	1000
Exact	0.496	0.193	$8.79\ 10^{-5}$	$5.67\ 10^{-9}$	$1.84\ 10^{-84}$
Large q	0.538	0.205	$9.04\ 10^{-5}$	$5.78\ 10^{-9}$	$1.85\ 10^{-84}$

Extension to bounded unary height?
(work in progress)

Further questions

- Bounded unary height: some values of k are special
  $\Rightarrow$ why? Combinatorial explanation?
- Asymptotic enumeration of $BCK(p)$ terms?
- Better bounds on asymptotic enumeration of unrestricted lambda-terms?
- Average unary height?
- Average arity of a unary node?
- Global height of a term? width? profile?
- u.s.w.