Statistical properties of lambda terms

Maciej Bendkowski   Olivier Bodini
Sergey Dovgal

CLA, Jussieu, Paris, 24/05/2018

Chapter 1. *Historical overview*

What is a closed lambda term?

Expression such as

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

Expression such as

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

- $\lambda -$ abstraction

Expression such as

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

- λ – abstraction
- @ – application

Expression such as

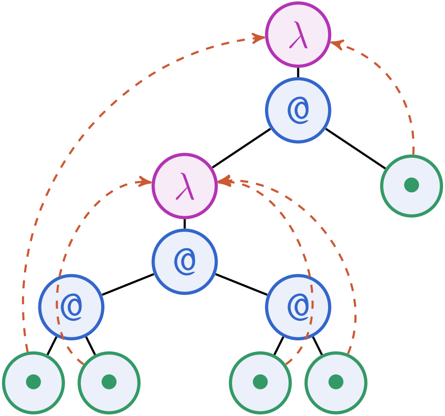
$$\lambda x.\lambda y.(\lambda z.(\lambda x.zx)(\lambda y.zy))(xy)$$

- λ — abstraction
- @ — application
- x, y, z — variables

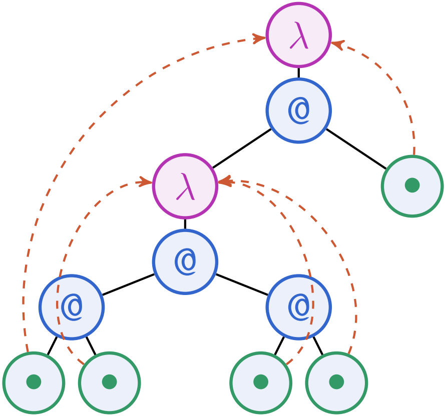
also lambda terms are..

Unary-binary trees with links

Unary-binary trees with links



Unary-binary trees with links



- λ are *abstractions*
- @ are *applications*
- • are *variables*

Local summary

Closed lambda terms are:

Local summary

Closed lambda terms are:

- Expressions such as $\lambda x.\lambda y.(\lambda z.(\lambda x.zx)(\lambda y.zy))(xy)$
- Unary-binary trees with links
- TODO: What are plain lambda terms?
- TODO: What is the difference between closed and plain lambda terms?

Local summary

Closed lambda terms are:

- Expressions such as $\lambda x.\lambda y.(\lambda z.(\lambda x.zx)(\lambda y.zy))(xy)$
- Unary-binary trees with links
- TODO: What are plain lambda terms?
- TODO: What is the difference between closed and plain lambda terms?

Local summary

Closed lambda terms are:

- Expressions such as $\lambda x.\lambda y.(\lambda z.(\lambda x.zx)(\lambda y.zy))(xy)$
- Unary-binary trees with links
- TODO: What are plain lambda terms?
- TODO: What is the difference between closed and plain lambda terms?

Local summary

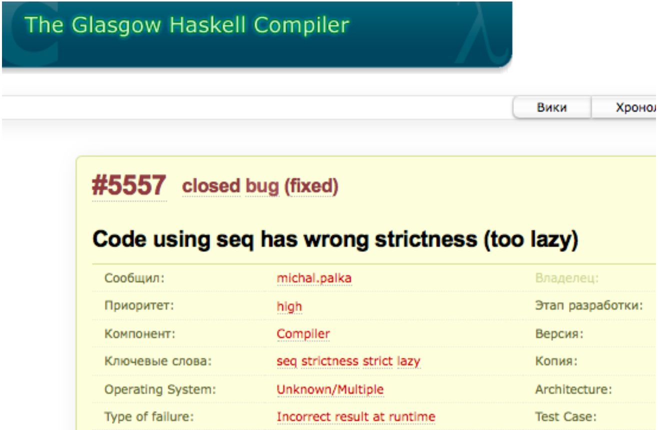
Closed lambda terms are:

- Expressions such as $\lambda x.\lambda y.(\lambda z.(\lambda x.zx)(\lambda y.zy))(xy)$
- Unary-binary trees with links
- TODO: What are plain lambda terms?
- TODO: What is the difference between closed and plain lambda terms?

Motivation

First application. Software testing techniques.

First application. Software testing techniques.



Second motivation. Relation between (linear) lambda terms and (trivalent) maps.

Third motivation. Development of analytic combinatorics.

Motivation

Summary

- Software testing techniques
- Lambda terms vs. maps
- Development of analytic combinatorics

Plain and closed lambda terms

- **Closed terms**: unary-binary trees with links between variables and abstractions
- **Plain terms**: unary-binary trees with some variables unlinked

Example.

Example.

- $\lambda z.(\lambda y.zy)$ is a **closed** term

Example.

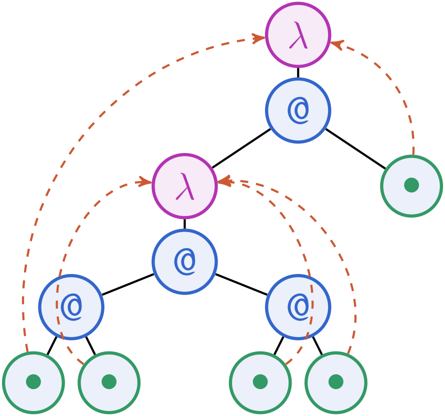
- $\lambda z.(\lambda y.zy)$ is a **closed** term
- $\lambda y.zy$ is a **plain** term

Example.

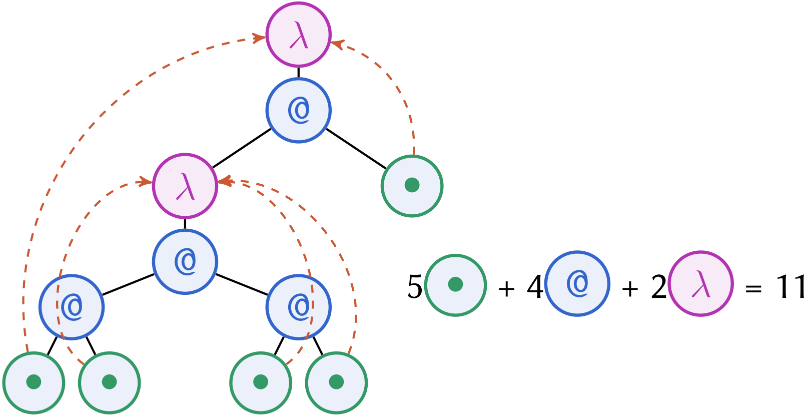
- $\lambda z.(\lambda y.zy)$ is a **closed** term
- $\lambda y.zy$ is a **plain** term
- $(\lambda x.x)(\lambda y.xy)$?

Size notion of a lambda term

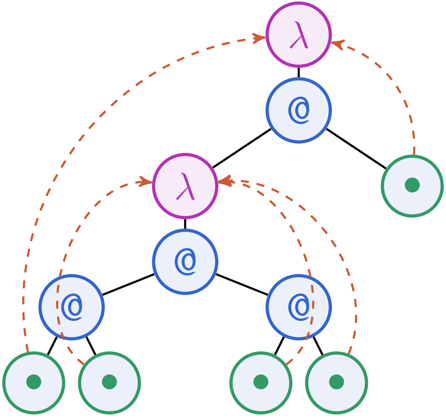
What is the size of



First approach. Total number of nodes.

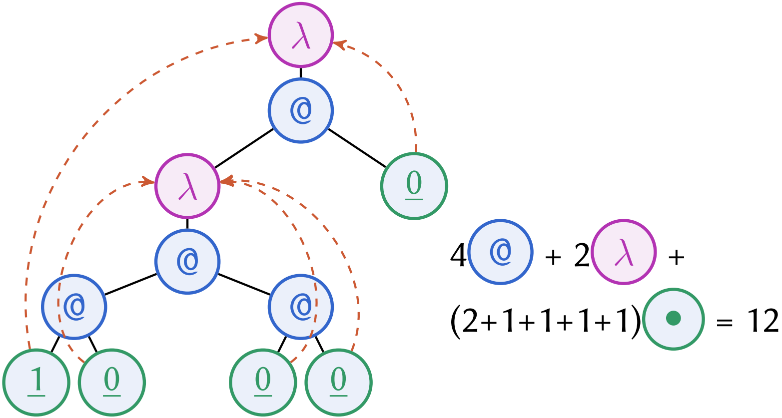


Second approach. Only abstractions and applications.

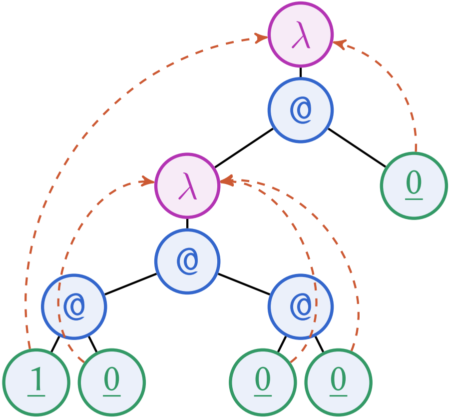


4 @ + 2 λ = 6

Third approach. Natural counting.



Fourth approach. Generalised natural counting.



4a @ +2b λ +5c • +

(1+0+0+0+0)d •

= 4a + 2b + 5c + d

Size notions, recap

- Variable size = 1
- Variable size = 0
- Variable size $\geqslant 1$ (Natural counting)
- Generalised natural counting

Size notions, recap

- Variable size = 1
- Variable size = 0
- Variable size $\geqslant 1$ (Natural counting)
- Generalised natural counting

Size notions, recap

- Variable size = 1
- Variable size = 0
- Variable size $\geqslant 1$ (Natural counting)
- Generalised natural counting

Size notions, recap

- Variable size = 1
- Variable size = 0
- Variable size $\geqslant 1$ (Natural counting)
- Generalised natural counting

In-detail analysis of size notions

1. Variable size equal 1

Variable size equal 1

- [Bodini, Gardy, Gittenberger, Jacquot '11+]
- Linear closed lambda terms
- Upper and lower bounds for asymptotics of all closed terms
- New bijections
- Continued [Zeilberger, Giorgetti + '15+]
- :(Formal power series is divergent

Variable size equal 1

- [Bodini, Gardy, Gittenberger, Jacquot '11+]
- Linear closed lambda terms
- Upper and lower bounds for asymptotics of all closed terms
- New bijections
- Continued [Zeilberger, Giorgetti + '15+]
- :(Formal power series is divergent

Variable size equal 1

- [Bodini, Gardy, Gittenberger, Jacquot '11+]
- Linear closed lambda terms
- Upper and lower bounds for asymptotics of all closed terms
- New bijections
- Continued [Zeilberger, Giorgetti + '15+]
- :(Formal power series is divergent

Variable size equal 1

- [Bodini, Gardy, Gittenberger, Jacquot ’11+]
- Linear closed lambda terms
- Upper and lower bounds for asymptotics of all closed terms
- New bijections
- Continued [Zeilberger, Giorgetti + ’15+]
- :(Formal power series is divergent

Variable size equal 1

- [Bodini, Gardy, Gittenberger, Jacquot '11+]
- Linear closed lambda terms
- Upper and lower bounds for asymptotics of all closed terms
- New bijections
- Continued [Zeilberger, Giorgetti + '15+]
- :(Formal power series is divergent

Variable size equal 1

- [Bodini, Gardy, Gittenberger, Jacquot ’11+]
- Linear closed lambda terms
- Upper and lower bounds for asymptotics of all closed terms
- New bijections
- Continued [Zeilberger, Giorgetti + ’15+]
- :(Formal power series is divergent

2. Variable size equal 0

Variable size equal 0

- [Grygiel, Lescanne ’12]
- Enumeration and random generation
- [David, Grygiel, Kozik, Raffalli, Theyssier, Zaionc ’13]
- $\Theta(\sqrt{n/\log n})$ head abstractions
- :(Formal power series is divergent

Variable size equal 0

- [Grygiel, Lescanne '12]
- Enumeration and random generation
- [David, Grygiel, Kozik, Raffalli, Theyssier, Zaionc '13]
- $\Theta(\sqrt{n/\log n})$ head abstractions
- :(Formal power series is divergent

Variable size equal 0

- [Grygiel, Lescanne '12]
- Enumeration and random generation
- [David, Grygiel, Kozik, Raffalli, Theyssier, Zaionc '13]
- $\Theta(\sqrt{n/\log n})$ head abstractions
- :(Formal power series is divergent

Variable size equal 0

- [Grygiel, Lescanne '12]
- Enumeration and random generation
- [David, Grygiel, Kozik, Raffalli, Theyssier, Zaionc '13]
- $\Theta(\sqrt{n/\log n})$ head abstractions
- :(Formal power series is divergent

Variable size equal 0

- [Grygiel, Lescanne ’12]
- Enumeration and random generation
- [David, Grygiel, Kozik, Raffalli, Theyssier, Zaionc ’13]
- $\Theta(\sqrt{n/\log n})$ head abstractions
- :(Formal power series is divergent

3. Natural counting

Natural counting

- Variable size = unary distance to binding λ
- [Bendkowski, Grygiel, Lescanne, Zaionc '12]
- [Bodini, Gittenberger, Gołębiewski '16]
  Enumeration and asymptotics
- :) Formal power series is convergent
- Uniform generation by Boltzmann sampling

Natural counting

- Variable size = unary distance to binding λ
- [Bendkowski, Grygiel, Lescanne, Zaionc ’12]
- [Bodini, Gittenberger, Gołębiewski ’16]
  Enumeration and asymptotics
- :) Formal power series is convergent
- Uniform generation by Boltzmann sampling

Natural counting

- Variable size = unary distance to binding λ
- [Bendkowski, Grygiel, Lescanne, Zaionc '12]
- [Bodini, Gittenberger, Gołębiewski '16]
  Enumeration and asymptotics
- :) Formal power series is convergent
- Uniform generation by Boltzmann sampling

Natural counting

- Variable size = unary distance to binding λ
- [Bendkowski, Grygiel, Lescanne, Zaionc ’12]
- [Bodini, Gittenberger, Gołębiewski ’16]
  Enumeration and asymptotics
- :) Formal power series is convergent
- Uniform generation by Boltzmann sampling

Natural counting

- Variable size = unary distance to binding λ
- [Bendkowski, Grygiel, Lescanne, Zaionc '12]
- [Bodini, Gittenberger, Gołębiewski '16]
  Enumeration and asymptotics
- :) Formal power series is convergent
- Uniform generation by Boltzmann sampling

Why consider a simpler model?

Why consider a simpler model?

Main reason: advantages in random generation

Why consider a simpler model?

Marking variables imply control over expectations

Why consider a simpler model?

Another reason: parameter study

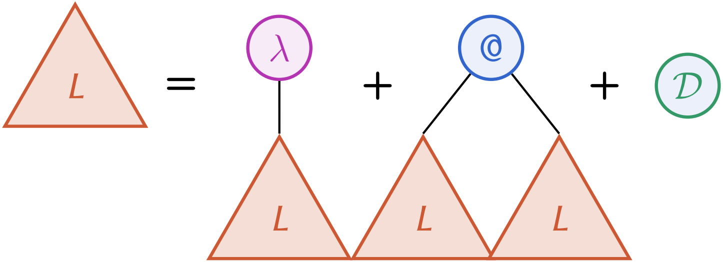
Why consider a simpler model?

Marking variables allow to study distributions

Chapter 2. *Statistical properties*

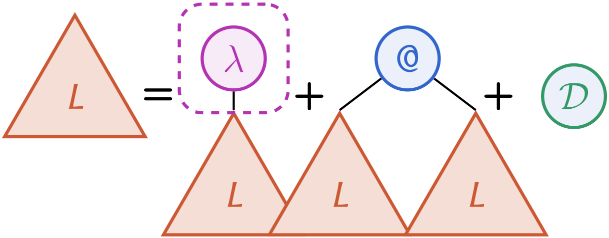
Plain lambda terms

Generating function for plain lambda terms



$$L(z) = zL(z) + zL^2(z) + \frac{z}{1-z}$$

Abstractions in plain terms?



$$L(z,u) = z{\color{purple}u}L(z,u) + zL(z,u) + \frac{z}{1-z}$$

$$L(z,u) \sim a(z,u) - b(z,u)\sqrt{1 - \frac{z}{\rho(u)}}$$

Multivariate Central Limit Theorem

Step 1. Extract coefficient of $L(z, u)$

$$L(z, u) \sim a(z, u) - b(z, u)\sqrt{1 - \frac{z}{\rho(u)}}$$

Step 2. Asymptotic behaviour of probability generating function

$$p_n(u) \sim A(u)B(u)^n$$

Step 3. Gaussian approximation from $A(u)B(u)^n$.

Applying multivariate CLT

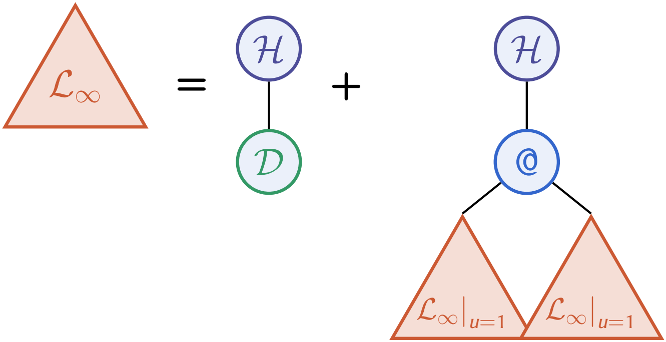
Theorem Joint Gaussian distribution for

- number of abstractions
- number of variables
- number of *redexes*

in plain lambda terms

Discrete distributions in plain terms

Head abstractions



$$L(z, u) = \frac{1}{1 - zu} \left(\frac{z}{1 - z} + zL(z, 1)^2 \right)$$

Theorem The following statistics follow discrete (geometric) limiting distributions

- The number of head abstractions
- Randomly chosed value of de Bruijn index

Chapter 3. *Infinite systems*

Drmota–Lalley–Woods theorem

Drmota–Lalley–Woods theorem

- Let $F(z)$ be a generating function
- Suppose it satisfies

$$F(z) = \Phi(F(z), z)$$

with Φ having combinatorial origin
- Then,

$$F(z) \sim a - b\sqrt{1 - \frac{z}{\rho}}$$

Drmota–Lalley–Woods theorem

- Let $F(z)$ be a generating function
- Suppose it satisfies

$$F(z) = \Phi(F(z), z)$$

with Φ having combinatorial origin

- Then,

$$F(z) \sim a - b\sqrt{1 - \frac{z}{\rho}}$$

Drmota–Lalley–Woods theorem

- Let $F(z)$ be a generating function
- Suppose it satisfies

$$F(z) = \Phi(F(z), z)$$

  with Φ having combinatorial origin
- Then,

$$F(z) \sim a - b\sqrt{1 - \frac{z}{\rho}}$$

What happens with infinite systems?

- [Drmota, Gittenberger, Morgenbesser '12]
- *If* Jacobian is a sum of identity matrix and a compact operator
- *And* specification is strongly connected
- *Then* Infinite-dimensional version holds

Closed lambda terms satisfy an infinite system

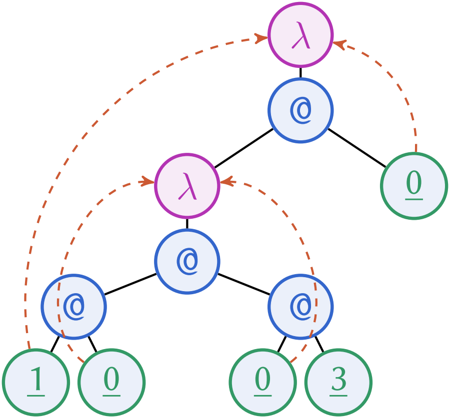
$$
\begin{aligned}
L_0(z) &= zL_1(z) + zL_0(z)^2 \quad , \\
L_1(z) &= zL_2(z) + zL_1(z)^2 + z \quad , \\
L_2(z) &= zL_3(z) + zL_2(z)^2 + z + z^2 \quad , \\
&\dots \\
L_\infty(z) &= zL_\infty(z) + zL_\infty(z) + \frac{z}{1-z} \quad .
\end{aligned}
$$

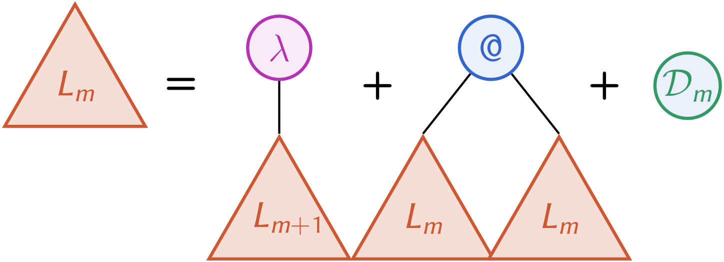
Why?

Adding m abstractions on the top of L_m makes the term closed

Example.

Example.





$$L_m(z) = zL_{m+1}(z) + zL_m^2(z) + z\frac{1 - z^m}{1 - z}$$

However

This system

$$
\begin{aligned}
L_0(z) &= zL_1(z) + zL_0(z)^2 \quad , \\
L_1(z) &= zL_2(z) + zL_1(z)^2 + z \quad , \\
L_2(z) &= zL_3(z) + zL_2(z)^2 + z + z^2 \quad , \\
&\cdots \\
L_\infty(z) &= zL_\infty(z) + zL_\infty(z) + \frac{z}{1-z} \quad .
\end{aligned}
$$

Doesn’t satisfy infinite DLW theorem

New “master theorem”

Master theorem

Motivation

Master theorem

Motivation

$$L_m(z) = zL_{m+1}(z) + zL_m(z)^2 + \underbrace{z\frac{1-z^m}{1-z}}_{D_m} ,$$

$$L_\infty(z) = zL_\infty(z) + zL_\infty(z)^2 + \underbrace{z\frac{1}{1-z}}_{D_\infty} .$$

The difference $D_\infty - D_m$ is exponentially small

Master theorem

More generally.

$$L_m(z) = K_m(L_m, L_{m+1}, z, u) \;\;,$$
$$L_\infty(z) = K_\infty(L_\infty, L_\infty, z, u) \;\;,$$

$K_\infty - K_m$ is exponentially small specified at L_∞
L_m, K_m, u are vectors

Master theorem

Assumptions

1. The limiting system has Puiseux expansion
2. Limiting system dominates *m*-th system
3. The difference is exponentially small

Statement

- Each L_m also has Puiseux expansion*

Each L_m also has Puiseux expansion*

Two possible variants:

1. (Weak)
$$[z^n]L_m(z,u) \sim [z^n]\left(a(u) - b(u)\sqrt{1 - \frac{z}{\rho(u)}}\right)$$
2. (Strong)
The analytic continuation of $L_m(z,u)$ in a *delta-domain* satisfies

$$L_m(z,u) \sim a(u) - b(u)\sqrt{1 - \frac{z}{\rho(u)}}$$

Lambda terms Infinite systems

Applications of “master theorem”

Theorem Joint Gaussian distribution for

- number of abstractions
- number of variables
- number of successors
- number of *redexes*

in closed lambda terms

Theorem Discrete distributions for

- number of head abstractions
- randomly chosed index value
- redex search time
- free variables
- missing top abstractions

in <u>closed</u> lambda terms

Theorem Rayleigh distributions for

- unary height profile
- natural height profile

in plain and closed lambda terms for variables, abstractions and applications

Summary for statistics

- Normal limit laws
- New discrete limit laws
- Rayleigh limit laws

Summary for statistics

- Normal limit laws
- New discrete limit laws
- Rayleigh limit laws

Summary for statistics

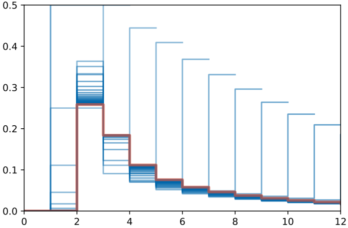
- Normal limit laws
- New discrete limit laws
- Rayleigh limit laws

Grand summary for statistics

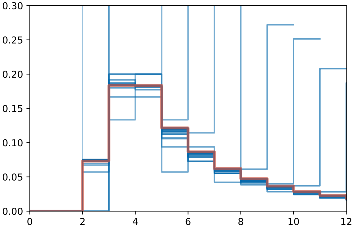
Parameter	Mean, $\sim$ plain	Mean, $\sim$ closed	Distribution plain	Distribution closed
Variables	$0.307n$		Normal	
Abstractions	$0.258n$		Normal	
Successors	$0.129n$		Normal	
Redexes	$0.091n$		Normal	
Index value	0.420		Geometric	
Redex search time	6.222	6.054	Discrete	Discrete
Head abstractions	0.420	1.447	Geometric	Discrete
m-openness	2.019	0	Discrete	trivial
Free variables	5.722	0	Discrete	trivial
Unary height profile	$0.122\sqrt{n}$		Rayleigh	
Natural height profile	$0.412\sqrt{n}$		Rayleigh	

Epilogue. *Open problems and experiments*

Redex search time distribution

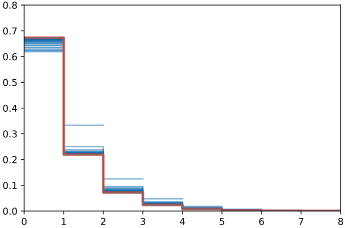


plain terms

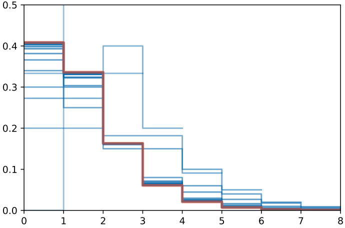


closed terms

Head abstractions

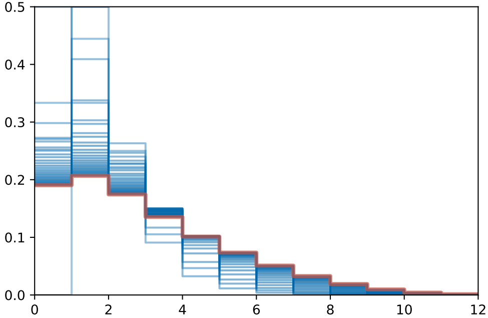


plain terms

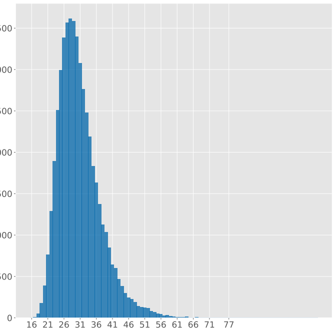


closed terms

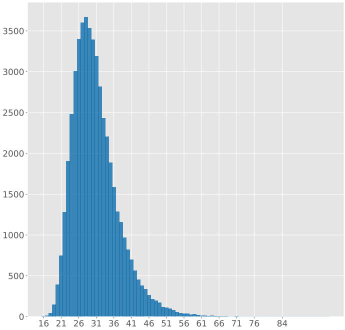
Free variables



Maximum number of variables bound to a single abstraction

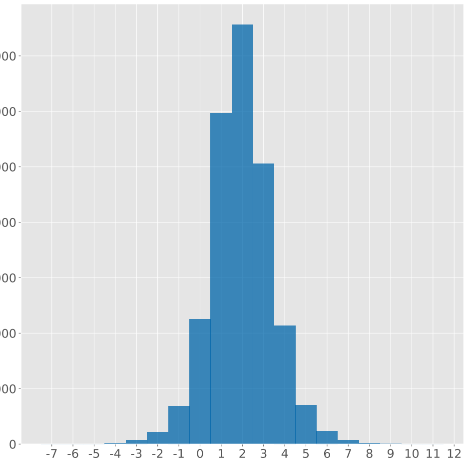


plain terms

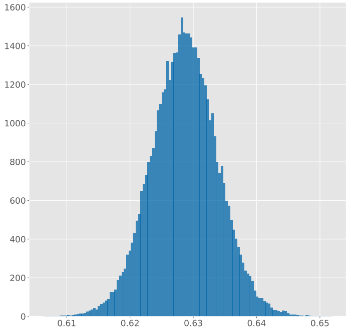


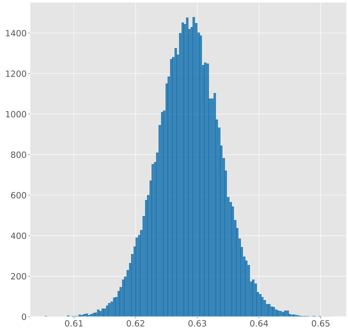
closed terms

Generalised m-openness



Number of binding abstractions





plain terms

closed terms

Open questions

- Maximum number of variables bound to a single abstraction
- Number of binding abstractions
- Generalised m-openness
- Number of BCI, BCK terms in natural size notion
- Average shape of a random lambda term after k β-reductions?

Open questions

- Maximum number of variables bound to a single abstraction
- Number of binding abstractions
- Generalised *m*-openness
- Number of BCI, BCK terms in natural size notion
- Average shape of a random lambda term after k β-reductions?

Open questions

- Maximum number of variables bound to a single abstraction
- Number of binding abstractions
- Generalised *m*-openness
- Number of BCI, BCK terms in natural size notion
- Average shape of a random lambda term after k β-reductions?

Open questions

- Maximum number of variables bound to a single abstraction
- Number of binding abstractions
- Generalised *m*-openness
- Number of BCI, BCK terms in natural size notion
- Average shape of a random lambda term after k β-reductions?

Open questions

- Maximum number of variables bound to a single abstraction
- Number of binding abstractions
- Generalised *m*-openness
- Number of BCI, BCK terms in natural size notion
- Average shape of a random lambda term after k β-reductions?

Conclusions

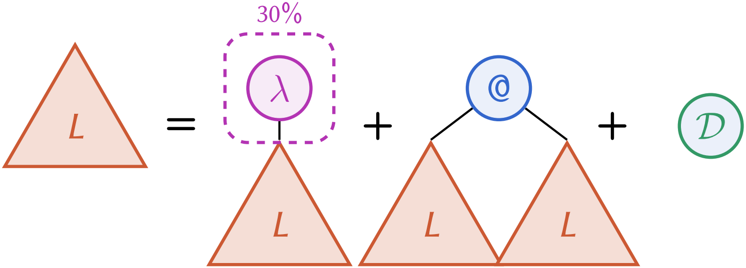
First.

First. Many tractable parameters in closed lambda terms in de Bruijn size notion

Second.

Second. Efficient multiparametric Boltzmann sampling, tweaking the desired parameters

Second. Efficient multiparametric Boltzmann sampling, tweaking the desired parameters



Third.

Third. Infinite systems of algebraic equations*

Third. Infinite systems of algebraic equations*

Conjecture Premises of DLW Theorem for infinite strongly connected systems are not sufficient.

Thank you !