Distribution of parameters in certain fragments of the linear and planar λ-calculus

Alexandros Singh, presenting joint work with Olivier Bodini and Noam Zeilberger.

October 6, 2020

Laboratoire Informatique de Paris Nord - CALIN Seminar

What is the λ-calculus?

- A *universal* formal system for expressing computation.
- Its terms are formed using the following grammar:
  - A variable is a valid term.
  - If x a variable and t is a valid term, then so is $(\lambda x.t)$.
  - If s and t are valid terms, then so is $(s\ t)$.
- The λ calculus also provides us with tools to transform terms, including the operation of β-reduction:

$$((\lambda x.t)\ s) \xrightarrow{\beta} t[x := s]$$

Some examples of terms:

$$(\lambda x.(xx))(\lambda x.(xx))$$

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

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

What is the λ-calculus?

- A *universal* formal system for expressing computation.
- Its terms are formed using the following grammar:
  - A variable is a valid term.
  - If x a variable and t is a valid term, then so is $(\lambda x.t)$.
  - If s and t are valid terms, then so is $(s\ t)$.
- The λ calculus also provides us with tools to transform terms, including the operation of β-reduction:

$$((\lambda x.t)\ s) \xrightarrow{\beta} t[x := s]$$

Some examples of terms:

$$(\lambda x.(xx))(\lambda x.(xx))$$

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

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

What is the λ-calculus?

- A *universal* formal system for expressing computation.
- Its terms are formed using the following grammar:
  - A variable is a valid term.
  - If x a variable and t is a valid term, then so is $(\lambda x.t)$.
  - If s and t are valid terms, then so is $(s\ t)$.
- The λ calculus also provides us with tools to transform terms, including the operation of β-reduction:

$$((\lambda x.t)\ s) \xrightarrow{\beta} t[x := s]$$

Some examples of terms:

$$(\lambda x.(xx))(\lambda x.(xx))$$

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

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

Combinatorics of the λ-calculus

- General terms are quite complicated. Growth is super-exponential, generating functions are not analytic. [1]
  Asymptotic number of general terms still (?) unresolved!
- We focus on *linear* terms: bound variables must appear exactly once: $\lambda x.(x\ x)$, $\lambda x.\lambda y.(a\ (y\ x))$.
- We also consider *planar* terms: bound variables must appear in the order they are introduced: $\lambda x.\lambda y.(y\ x)$, $\lambda x.\lambda y.(a\ (x\ y))$.

[1] For the notion of term size given recursively by:
$|var| = 1$, $|(s\ t)| = |s| + |t| + 1$, $|\lambda x.t| = |t| + 1$.

Combinatorics of the λ-calculus

- General terms are quite complicated. Growth is super-exponential, generating functions are not analytic. [1] Asymptotic number of general terms still (?) unresolved!
- We focus on *linear* terms: bound variables must appear exactly once: $\lambda x.(x\ x)$, $\lambda x.\lambda y.(a\ (y\ x))$.
- We also consider *planar* terms: bound variables must appear in the order they are introduced: $\lambda x.\lambda y.(y\ x)$, $\lambda x.\lambda y.(a\ (x\ y))$.

[1] For the notion of term size given recursively by: $|var| = 1$, $|(s\ t)| = |s| + |t| + 1$, $|\lambda x.t| = |t| + 1$.

Combinatorics of the λ-calculus

- General terms are quite complicated. Growth is super-exponential, generating functions are not analytic. [1] Asymptotic number of general terms still (?) unresolved!
- We focus on *linear* terms: bound variables must appear exactly once: $\lambda x.(x\ x)$, $\lambda x.\lambda y.(a\ (y\ x))$.
- We also consider *planar* terms: bound variables must appear in the order they are introduced: $\lambda x.\lambda y.(y\ x)$, $\lambda x.\lambda y.(a\ (x\ y))$.

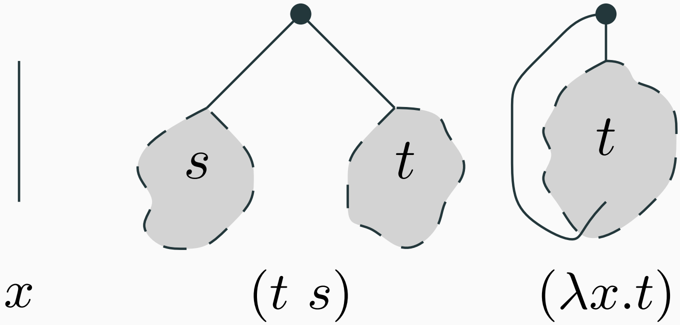
[1] For the notion of term size given recursively by: $|var| = 1$, $|(s\ t)| = |s| + |t| + 1$, $|\lambda x.t| = |t| + 1$.

The λ-calculus and maps

- Maps: graphs embedded in an oriented surface without boundary.
- Closed linear terms are combinatorially intriguing: they correspond to rooted connected trivalent maps! [1, 2]
  Closed planar terms correspond to *planar* such maps. Open terms allow for univalent vertices too.

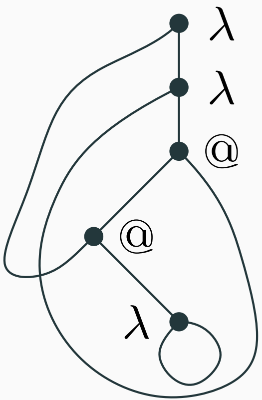
The λ-calculus and maps

- Maps: graphs embedded in an oriented surface without boundary.
- Closed linear terms are combinatorially intriguing: they correspond to rooted connected trivalent maps! [1, 2]
  Closed planar terms correspond to *planar* such maps. Open terms allow for univalent vertices too.



An example of a term and its corresponding map

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



Where λ annotates abstractions and @ applications.

Purpose of this work

- How do ”typical” (random, of large size) linear and planar terms behave?
- How many free variables do they have? How often is a typical term an abstraction?
- Using tools from analytic combinatorics to obtain parameter distributions.

Purpose of this work

- How do ”typical” (random, of large size) linear and planar terms behave?
- How many free variables do they have? How often is a typical term an abstraction?
- Using tools from analytic combinatorics to obtain parameter distributions.

Purpose of this work

- How do ”typical” (random, of large size) linear and planar terms behave?
- How many free variables do they have? How often is a typical term an abstraction?
- Using tools from analytic combinatorics to obtain parameter distributions.

In this talk

We'll sketch the following results:

Linear λ-Terms
(Differentially Algebraic, Divergent)

- Limit distribution of free variables
- Limit distribution of identity-subterms in closed terms.
- Limit distribution of closed subterms in closed terms.
- Probability that term is an abstraction.

Planar λ-Terms
(Algebraic, Analytic)

- Limit distribution of free variables for regular and bridgeless terms.
- Probability that regular or bridgeless open term is an abstraction.

Free variables in closed linear terms

- Free variables are those not bound by an abstraction. For example: $\lambda x.(a\ x)$

Proposition

The limit distribution of free variables in linear λ-terms of size n is Gaussian with mean and variance $\mu = \sigma^2 \sim \sqrt[3]{n}$.

Starting point (follows from definition of combinatorial maps):

$$L(z^2, u) = uz^2 + z^4 + z^5 \frac{\partial}{\partial z} \left(\ln \left(\exp(z^2/2) \odot \exp(z^3/3 + uz) \right) \right)$$

where L counts open linear λ-terms with u tagging free variables.

Free variables in closed linear terms

- Free variables are those not bound by an abstraction. For example: $\lambda x.(a\ x)$

Proposition

The limit distribution of free variables in linear λ-terms of size n is Gaussian with mean and variance $\mu = \sigma^2 \sim \sqrt[3]{n}$.

Starting point (follows from definition of combinatorial maps):

$$L(z^2, u) = uz^2 + z^4 + z^5 \frac{\partial}{\partial z} \left(\ln \left(\exp(z^2/2) \odot \exp(z^3/3 + uz) \right) \right)$$

where L counts open linear λ-terms with u tagging free variables.

Proof Sketch

Saddle-point analysis of Hadamard product yields:

$$[z^n]\exp\left(z^3/3+uz\right) = \left(\frac{1}{6}\frac{\sqrt{2}\sqrt{3}\,n^{-\frac{1}{2}}}{\sqrt{\pi}} - \frac{1}{36}\frac{\sqrt{2}\sqrt{3}u^2n^{-\frac{5}{6}}}{\sqrt{\pi}} + O\left(n^{-\frac{7}{6}}\right)\right) e^{un^{1/3}+n/3}n^{-n/3}$$

$$[z^n]\exp\left(z^2/2\right) \sim \frac{1}{2}\,\frac{e^{1/2+n/2}}{\left(\sqrt{1+n}\right)^{1+n}\sqrt{\pi}} - \frac{1}{2}\,\frac{e^{1/2+n/2}}{\left(-\sqrt{1+n}\right)^{1+n}\sqrt{\pi}}$$

While an application of Bender's theorem [3, Theorem 1] gives

$$2[z^n]\frac{d}{dz}\ln(A(z^{1/2},u)) = n\left([z^n]h(z,u) - \frac{1}{2}[z^{n-2}]h(z,u)\right) + O\left([z^{n-4}]h(z,u)\right)$$

for $A(x,u) = \exp(z^2/2) \odot \exp(z^3/3+uz)$

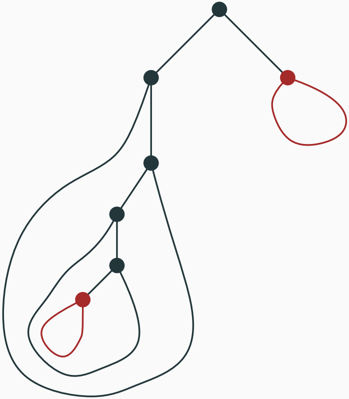
Distribution of identity-subterms in closed linear terms

- Identity terms: terms which are α-equivalent to $\lambda x.x$. For example: $\lambda x.(x\ (\lambda y.y))$.
- They appear as loops in the corresponding map.

Distribution of identity-subterms in closed linear terms

- Identity terms: terms which are α-equivalent to $\lambda x.x$. For example: $\lambda x.(x\ (\lambda y.y))$.
- They appear as loops in the corresponding map.

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



Distribution of identity-subterms in closed linear terms

Proposition

The limit distribution of identity-subterms in closed linear λ-terms is Poisson of parameter $\lambda = 1$.

Proof Sketch: Use moment pumping on

$$G = (u-1)z^2 + zG^2 + \frac{\partial}{\partial u} G$$

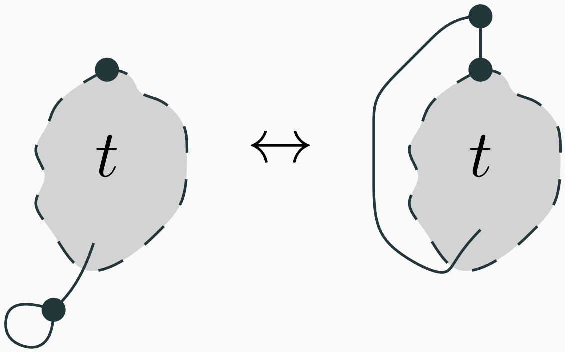
where G counts closed linear terms with u tagging identity-subterms.

Distribution of identity-subterms in closed linear terms

Justification for

$$G = (u-1)z^2 + zG^2 + \frac{\partial}{\partial u}G$$

Terms are either identity-terms, applications, or



Distribution of identity-subterms in closed linear terms

For the pumping, note that the k-th derivative of the eq. may be written as

$$\frac{\partial^k}{\partial u^k} G - S - 2z\, G\, \frac{\partial^k}{\partial u^k} G = \frac{\partial^{k+1}}{\partial u^{k+1}} G$$

with S, depending on the parity of k, being as follows

$$\sum_{l=1}^{\lfloor \frac{k}{2} \rfloor} 2z \binom{k}{l} \frac{\partial^l}{\partial u^l} G \frac{\partial^{k-l}}{\partial u^{k-l}} G, \text{ for odd } k$$

$$\sum_{l=1}^{\lfloor \frac{k}{2} \rfloor - 1} 2z \binom{k}{l} \frac{\partial^l}{\partial u^l} G \frac{\partial^{k-l}}{\partial u^{k-l}} G + z \binom{k}{\lfloor \frac{k}{2} \rfloor} \left(\frac{\partial^{\lfloor \frac{k}{2} \rfloor}}{\partial u^{\lfloor \frac{k}{2} \rfloor}} G \right)^2, \text{ for even } k$$

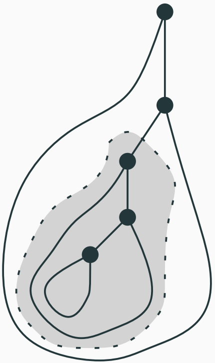
Distribution of closed subterms in closed linear terms

- Closed subterms: subterms having no free variables.
- Coresponding to non-root-containing connected components resulting from the deletion of some bridge in the respective map.

Distribution of closed subterms in closed linear terms

- Closed subterms: subterms having no free variables.
- Coresponding to non-root-containing connected components resulting from the deletion of some bridge in the respective map.

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



Distribution of closed subterms in closed linear terms

Proposition

The limit distribution of closed subterms in closed linear λ-terms is Poisson of parameter $\lambda = 1$.

Proof Sketch: Use moment pumping on

$$\frac{\partial W}{\partial v} = \frac{-(zv^2W^2 + z^2 - W)W}{zv^2(v-1)W^2 + (1-v)W + vz^2}$$

where W counts closed linear terms with v tagging identity-subterms.

Probability that a closed linear term is an abstraction

Proposition

Asymptotically almost surely a random closed linear λ-term is an abstraction.

Proof Sketch:

It can be shown that $[z^n]L_c \sim k \cdot 6^n \cdot n!$ for some constant k. Compare the coefficients of L_c and $2z^4 \frac{\partial}{\partial z} L_c$ in

$$L_c = z^2 + zL_c^2 + 2z^4 \frac{\partial}{\partial z} L_c.$$

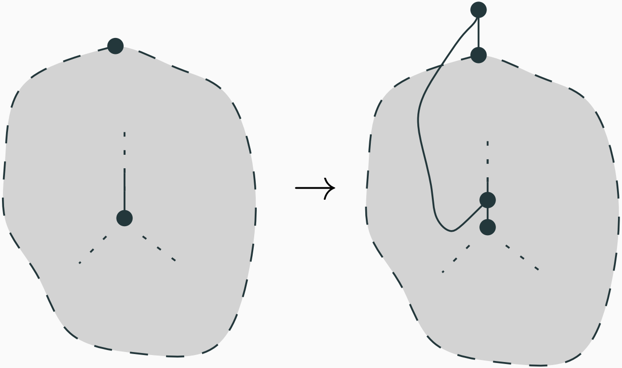
where L_c enumerates closed linear λ-terms.

Distribution of identity-subterms in closed linear terms

Justification for

$$G = L_c = z^2 + zL_c^2 + 2z^4 \frac{\partial}{\partial z} L_c.$$

Terms are either identity-terms, applications, or



Distribution of free variables in planar and bridgeless planar terms

Proposition

The limit distribution of free variables in planar λ-terms of size n is Gaussian with mean $\mu = \frac{n}{8}$ and variance $\sigma^2 = \frac{9n}{32}$.

Proposition

The limit distribution of free variables in bridgeless planar λ-terms of size n is Gaussian with mean $\mu = \frac{n}{5}$ and variance $\sigma^2 = \frac{9n}{25}$.

Distribution of free variables in planar and bridgeless planar terms

Both results follow similar steps.

Our starting points are the following two equations

$$P(z,u) = uz + zQ(z,u)^2 + \frac{z(P(z,u) - P(z,0))}{u}$$
$$Q(z,u) = uz + zQ(z,u)^2 + \frac{z\left(Q(z,u) - u[u^1]Q(z,u)\right)}{u}$$

with P and Q counting planar and bridgless planar terms respectively and u tagging free variables.

Sketch: use elimination and the quadratic method to obtain closed form solutions. Proceed by applying, [4, Proposition IX.17].

Distribution of free variables in planar and bridgeless planar terms

$$
\begin{aligned}
Q(z,u) &= 1/2\,z^{-1} - 1/2\,u^{-1} \\
&+ \frac{1}{2}\,\frac{1}{uz}\left(\frac{1}{3}\,u^{2}\sqrt[3]{-1458\,z^{6} + 6\,\sqrt{3}\sqrt{19683\,z^{8} - 4374\,z^{5} + 324\,z^{2} - 8\,z^{-1}}z^{2} - 270\,z^{3} + 1}\right. \\
&+ 36\,u^{2}z^{3}\,\frac{1}{\sqrt[3]{-1458\,z^{6} + 6\,\sqrt{3}\sqrt{19683\,z^{8} - 4374\,z^{5} + 324\,z^{2} - 8\,z^{-1}}z^{2} - 270\,z^{3} + 1}} \\
&+ \frac{1}{3}\,u^{2}\,\frac{1}{\sqrt[3]{-1458\,z^{6} + 6\,\sqrt{3}\sqrt{19683\,z^{8} - 4374\,z^{5} + 324\,z^{2} - 8\,z^{-1}}z^{2} - 270\,z^{3} + 1}} \\
&\left.+ \frac{1}{3}\,u^{2} - 4\,u^{3}z^{2} - 2\,uz + z^{2}\right)^{1/2}.
\end{aligned}
$$

Distribution of free variables in planar and bridgeless planar terms

While $P(z) = A(z,u) + B(z,u) \cdot C(z,u)^{-1/2}$ with

$$A(z,u) = \frac{1}{2z} - \frac{1}{2u},\ B(z,u) = \frac{1}{2uz}$$

$$\begin{aligned}C(Z,u) = & -4\,u^3 z^2 \\ & + \frac{1}{48}\,\frac{u\sqrt[3]{1492992\,z^{12} + 8640\,z^6 + 96\,\sqrt{3}\sqrt{80621568\,z^{18} - 559872\,z^{12} + 1296\,z^6 - 1}z^3 - 1}}{z^2} \\ & + 72\,\frac{uz^4}{\sqrt[3]{1492992\,z^{12} + 8640\,z^6 + 96\,\sqrt{3}\sqrt{80621568\,z^{18} - 559872\,z^{12} + 1296\,z^6 - 1}z^3 - 1}} \\ & + \frac{1}{48}\,\frac{u}{z^2\sqrt[3]{1492992\,z^{12} + 8640\,z^6 + 96\,\sqrt{3}\sqrt{80621568\,z^{18} - 559872\,z^{12} + 1296\,z^6 - 1}z^3 - 1}} \\ & - \frac{1}{48}\,\frac{u}{z^2} + u^2 + z^2\end{aligned}$$

Probability that an open planar or bridgeless planar term is an abstraction

Proposition

Asymptotically, the probability that an random open planar (bridgeless planar) term is an abstraction is $\rho_P = \frac{\sqrt{2}}{4}$ $(\rho_{PB} = \frac{2}{5})$.

Proof Sketch: Estimate

$$\frac{[z^n]\ z(P(z,1) - P(z,0))}{[z^n]\ P(z,1)} \text{ and } \frac{[z^n]\ z(Q(z,1) - ([u^1]Q(z,u))|_{u=1})}{[z^n]\ Q(z,1)}$$

Both $P(z,0)$ and $[u^1]Q(z,u)|_{u=1}$ are analytic at the respective singularities ρ_P and ρ_{PB} of P and Q. Use the singular expansions of P, Q at the corresponding singularities to obtain the desired result.

Conclusions

- Clear distinctions between the divergent/differentially-algebraic case of linear terms and the algebraic one of planar terms.
- Need for: more tools to handle divergent combinatorial classes, algebraicity results for *closed* planar terms.
- Future directions: study of β-reduction, typing of linear terms.

Thank you!

Conclusions

- Clear distinctions between the divergent/differentially-algebraic case of linear terms and the algebraic one of planar terms.
- Need for: more tools to handle divergent combinatorial classes, algebraicity results for *closed* planar terms.
- Future directions: study of β-reduction, typing of linear terms.

Thank you!

Conclusions

- Clear distinctions between the divergent/differentially-algebraic case of linear terms and the algebraic one of planar terms.
- Need for: more tools to handle divergent combinatorial classes, algebraicity results for *closed* planar terms.
- Future directions: study of β-reduction, typing of linear terms.

Thank you!

Bibliography

- O. Bodini, D. Gardy, and A. Jacquot, “Asymptotics and random sampling for BCI and BCK lambda terms,” *Theoretical Computer Science*, vol. 502, pp. 227–238, 2013.
- N. Zeilberger, “Linear lambda terms as invariants of rooted trivalent maps,” *Journal of Functional Programming*, vol. 26, 2016.
- E. A. Bender, “An asymptotic expansion for the coefficients of some formal power series,” *Journal of the London Mathematical Society*, vol. 2, no. 3, pp. 451–458, 1975.
- P. Flajolet and R. Sedgewick, *Analytic combinatorics*. cambridge University press, 2009.