Distribution of Parameters in Certain Fragments of the Linear and Planar λ-calculus

Olivier Bodini Alexandros Singh Noam Zeilberger

September 3, 2020

1 Introduction

Structural properties of large random λ-terms may be gleaned by studying the asymptotic distributions of various parameters of interest, such as the number of free variables, abstractions and applications, and so on. Such properties have been studied for general λ-terms under various considerations such as different notions of size and various structural restrictions (for examples, see [1, 2, 3, 3, 4]).

Linear λ-terms, that is, terms where each variable occurs exactly once, form an interesting subsystem of λ-calculus with various combinatorial connections to much-studied classes of objects such as trivalent maps [5, 6].

The purpose of this work is to help shed some light on what the "typical" terms of certain fragments of the linear λ calculus look like. The fragments we deal with in this work may be roughly partitioned into two major categories, the algebraic and the differentially-algebraic one, according to the nature of the specifications they admit.

This qualitative difference in specification also manifests itself on the level of parameter distributions, where we observed markedly different typical behaviours for parameters such as free-variable distributions.

Following this pattern, we present our results in two sections, each devoted to one of the corresponding aforementioned categories.

2 Algebraic Classes

The results presented in this section concern the classes of *planar* terms, that is terms in which variables are used in the same order they are bound.

Interestingly, the planarity restriction leads to algebraicity of the generating functions, since the specifications are amenable to the kernel method and elimination techniques.

We study both general planar terms and an interesting subclass of them consisting of terms having no proper closed subterms, called *bridgeless*.

These classes exhibit moderate growth and as such are amenable to analysis using standard tools of analytic combinatorics.

For example, the following two results on the Gaussian distributions of free variables in open planar and bridgeless planar terms are typical of such algebraic classes.

Proposition 2.1. *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 2.2. *The limit distribution of free variables in bridgeless planar λ-terms of size n is Gaussian with mean $\mu = \frac{n}{5}$ and variance $\sigma^2 = \frac{9n}{25}$.*

The starting point for proving these propositions is the following two equations characterising our classes

$$P(z,u) = uz + zP(z,u)^2 + \frac{z(P(z,u) - P(z,0))}{u} \tag{1}$$

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

where $P(z,0)$ counts closed planar λ-terms and $[u^1]Q(z,u)$ counts bridgeless planar terms with one free variable. The well-behaved nature of both P and Q proves very useful in obtaining the following results too.

Proposition 2.3. *Asymptotically, the probability that a random open planar term is an abstraction is $\rho_P = \frac{\sqrt{2}}{4}$.*

Proposition 2.4. *Asymptotically, the probability that a random open bridgeless planar term is an abstraction is $\rho_{PB} = \frac{2}{5}$.*

3 Differentially-Algebraic Classes

The results presented in this section concern classes whose growth is factorial. Accordingly, the tools used to analyse these classes are of a markedly different nature than the ones used to analyse planar terms in the previous section. Much of the analysis presented here relies on coming up with various different specifications for the classes considered, each of which allows us to access different parameters of interest.

Proposition 3.1. *The limit distribution of free variables in linear λ-terms of size n is Gaussian with mean and variance $\mu = \sigma^2 \sim \sqrt[3]{n}$.*

Proposition 3.2. *The limit distribution of identity-subterms in closed linear λ-terms is Poisson of parameter $\lambda = 1$.*

Proposition 3.3. *The limit distribution of closed subterms in closed linear λ-terms is a shifted Poisson distribution of parameter $\lambda = 1$.*

For example, Proposition 3.3 implies that as $n \to \infty$, the probability that a random closed linear λ-term of size n contains no closed proper subterms approaches $1/e$. This further implies that bridgeless linear terms may be uniformly generated in expected linear time by rejection sampling, by adapting Algorithm 1 of [5].

The rapid growth of closed linear λ-terms leads to results such as the following.

Proposition 3.4. *Asymptotically almost surely a random closed linear λ-term is an abstraction.*

References

[1] M. Bendkowski, O. Bodini, and S. Dovgal, "Statistical properties of lambda terms," *The Electronic Journal of Combinatorics*, pp. P4–1, 2019.

[2] B. Gittenberger and I. Larcher, "On the number of variables in special classes of random lambda-terms," in *29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018)*, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

[3] K. Grygiel and I. Larcher, "Unary profile of lambda terms with restricted De Bruijn indices," 2019.

[4] B. Gittenberger and I. Larcher, "Distribution of variables in lambda-terms with restrictions on De Bruijn indices and De Bruijn levels," *The Electronic Journal of Combinatorics*, pp. P4–47, 2019.

[5] 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.

[6] N. Zeilberger, "Linear lambda terms as invariants of rooted trivalent maps," *Journal of Functional Programming*, vol. 26, 2016.