Asymptotic Properties of Combinatory Logic

Maciej Bendkowski, Katarzyna Grygiel$^{(\boxtimes)}$, and Marek Zaionc

Theoretical Computer Science Department, Faculty of Mathematics and Computer Science, Jagiellonian University, ul. Łojasiewicza 6, 30-348 Kraków, Poland
{bendkowski,grygiel,zaionc}@tcs.uj.edu.pl

Abstract. We present a quantitative analysis of random combinatory logic terms. Our main goal is to investigate likelihood of semantic properties of random combinators. We show that asymptotically almost all weakly normalizing terms are not strongly normalizing. Moreover, we present a proof that asymptotically almost all strongly normalizing terms are not in normal form. We also prove that asymptotically almost all normal forms in combinatory logic are not typeable.

Keywords: Models of computation · Combinatory logic · Asymptotic probability in logic

1 Introduction

Over the last decade quantitative aspects of logic have attracted increasing attention from researchers working on the border of combinatorics, logic, and computer science. Probabilistic methods used in the paper appear to be very powerful in computer science investigations. From a point of view of these methods we study typical objects chosen from a given set. In recent years we have investigated sets of syntactic objects of logical flavor in order to estimate likelihood of the fact that a randomly chosen syntactic object belongs to a given set. There is a long history of using this kind of asymptotic approach applied to logic and computability. Probability of truth of logical formulas has been investigated in several papers. For the purely implicational logic of one variable (and at the same time simply typed system), the likelihood of finding true formulas was computed by Moczurad, Tyszkiewicz, and Zaionc in [14]. The classical logic of one variable and two connectives of implication and negation was studied in Zaionc [20]; over the same language, the exact proportion between intuitionistic and classical logics was determined by Kostrzycka and Zaionc in [11].

Asymptotic id entity between classical and intuitionistic logic of implication has been proved in Fournier, Gardy, Genitrini, and Zaionc in [6]. Some variants involving expressions with other logical connectives have also been considered.

Genitrini and Kozik in [9] have studied asymptotic behavior of full propositional system. For two connectives again, the *and/or* case has already received much attention—see Lefmann and Savický [13], Chauvin, Flajolet, Gardy, and Gittenberger [2], Gardy and Woods [8], Woods [19] and Kozik [12]. Let us also mention the survey [7] of Gardy on probability distributions on Boolean functions induced by random Boolean expressions.

In [4] investigations of computational objects from lambda calculus and combinatory logic were started. It was shown that a randomly chosen λ-term is strongly normalizing. In the case of combinatory logic (the equivalent translation of the λ-calculus), the situation is exactly opposite—a random combinator does not strongly normalize. Since every strongly normalizing term (in both models) is weakly normalizing, the obtained results imply that a random lambda term satisfies the weak normalization property, however, they do not allow us to claim anything about weakly normalizing combinators. The counting problem for lambda terms is still a very hot open research subject. Some variants of lambda calculus have also been considered. Bodini, Gardy, Gittenberger and Jacquot in [15] studied enumeration of BCI lambda terms. John Tromp in [17], as well as Grygiel and Lescanne in [10], considered the enumeration problem in the so called binary lambda calculus.

The syntax of combinators is very simple, as the terms in question can be uniquely represented by finite binary planar trees whose leaves are labeled by constants. In contrast to lambda calculus terms, whose unusual tree representation makes the combinatorial analysis very difficult (see, e.g. [4]), the analysis of combinators satisfying a given syntactic property is usually simple. However, in the case of properties that are undecidable, the enumeration problems become hard or even impossible, as for any nonrecursive set it is impossible to find a finite pattern collection defining the whole set. For example, any nonrecursive set of combinators cannot be defined by a context-free grammar. Therefore the only possible approach to find asymptotic behavior of nonrecursive sets of combinators is to construct proper recursive subsets and proper recursive supersets.

In this paper we give a simple argument that the density of weakly normalizing combinators is neither zero nor one. Moreover, we present lower and upper bounds for the density in question. This result allows us to compare two basic nonrecursive sets of combinators, one being the subset of another: the set of all weakly normalizing combinators and its proper subset—the set of all strongly normalizing combinators. It turns out, that the set of strongly normalizing terms can be seen as a tiny fragment of the set of weakly normalizing combinators. In other words, we prove that the asymptotic probability of finding strongly normalizing terms chosen from the set of weakly normalizing ones is zero.

Another part of the paper is oriented toward terms in so called normal forms. At the same time we are interested in typeable terms in combinatory logic which form an important subclass motivated by programming languages. Both typeable terms (in the simple type system) and terms in normal forms form recursive sets of combinators. In the paper we present a result concerning typeable normal forms in the setting of all normal forms.
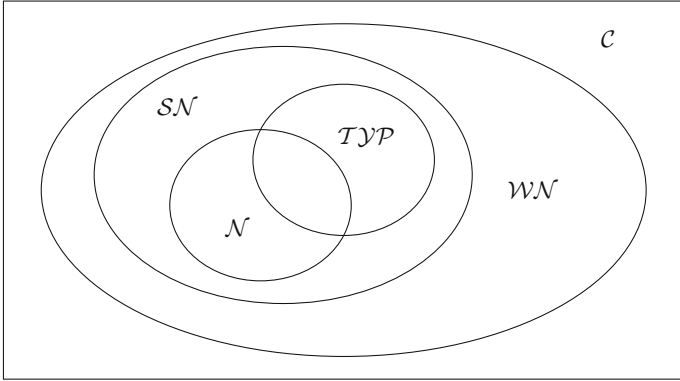
Fig. 1. Partition of combinatory logic

Figure 1 illustrates the topography of all major classes of combinatory logic terms used in this paper. $\mathcal{C}$ denotes the set of all combinators, $\mathcal{WN}$ stands for weakly normalizing terms, $\mathcal{SN}$ for strongly normalizing ones, $\mathcal{N}$ for normal forms and, finally, $\mathcal{TYP}$ stands for typeable terms.

2   Combinators

The set $\mathcal{C}$ of *combinators* is defined inductively as follows. Constants **K** and **S** are combinators. If M and N are combinators, then $(M\ N)$ is a combinator. Terms built as in the third case are called *applications*. Following standard notational conventions, we omit outermost parentheses and drop parentheses from left-associated terms, e.g., instead of $((MN)(PQ))$ we write $MN(PQ)$. We define a one-step reduction relation $\rightarrow$ on the set of combinators in the following way. Let P, Q, R be arbitrary combinators. Then

– $\mathbf{K}PQ \rightarrow P$,
– $\mathbf{S}PQR \rightarrow PR(QR)$,
– if $P \rightarrow Q$ then $PR \rightarrow QR$ and $RP \rightarrow RQ$.

Let P be a combinator. If there exists no combinator Q such that $P \rightarrow Q$, then P is said to be in *normal form*. If there exists a finite sequence of combinators $P_0, P_1, \ldots, P_k$ such that $P = P_0 \rightarrow P_1 \rightarrow \ldots \rightarrow P_k$ and P_k is in normal form, then P is *weakly normalizing*. If there does not exist an infinite sequence of combinators $P_0, P_1, \ldots$ such that $P = P_0 \rightarrow P_1 \rightarrow \ldots$, then we say that P is *strongly normalizing*. Of course, strong normalization implies weak normalization.

3   Densities of Sets of Combinators

With the set of all combinators $\mathcal{C}$ we associate the size function defined as the number of all applications occurring in a given combinator, i.e.,

$$|\mathbf{S}| = |\mathbf{K}| = 0 \quad \text{and} \quad |PQ| = 1 + |P| + |Q|.$$

Given a subset $\mathcal{X} \subseteq \mathcal{C}$ of combinators we define the *asymptotic density* $\mu(\mathcal{X})$ as

$$\mu(\mathcal{X}) = \lim_{n \to \infty} \frac{\#\{t \in \mathcal{X} : |t| = n\}}{\#\{t \in \mathcal{C} : |t| = n\}}$$

if the limit exists. The number $\mu(\mathcal{X})$ is an asymptotic probability of finding a combinator from the class $\mathcal{X}$ among all combinators. It can be immediately seen that the density μ is finitely additive, but not countably additive. Finally, we define:

$$\mu^-(\mathcal{X}) = \liminf_{n \to \infty} \frac{\#\{t \in \mathcal{X} : |t| = n\}}{\#\{t \in \mathcal{C} : |t| = n\}}$$

$$\mu^+(\mathcal{X}) = \limsup_{n \to \infty} \frac{\#\{t \in \mathcal{X} : |t| = n\}}{\#\{t \in \mathcal{C} : |t| = n\}}$$

These two numbers are well defined for any set $\mathcal{X}$ of combinators, even when the limiting ratio $\mu(\mathcal{X})$ is not known to exist. Given two classes of combinators $\mathcal{X}$ and $\mathcal{Y}$, assuming that $\mathcal{X}$ is a subset of $\mathcal{Y}$, we define *relative density* $\mu\left(\frac{\mathcal{X}}{\mathcal{Y}}\right)$ in the usual way by allowing:

$$\mu\left(\frac{\mathcal{X}}{\mathcal{Y}}\right) = \lim_{n \to \infty} \frac{\#\{t \in \mathcal{X} : |t| = n\}}{\#\{t \in \mathcal{Y} : |t| = n\}}$$

The relative $\mu^-\left(\frac{\mathcal{X}}{\mathcal{Y}}\right)$ and $\mu^+\left(\frac{\mathcal{X}}{\mathcal{Y}}\right)$ functions are defined in the very same way as in general case, i.e.,

$$\mu^-\left(\frac{\mathcal{X}}{\mathcal{Y}}\right) = \liminf_{n \to \infty} \frac{\#\{t \in \mathcal{X} : |t| = n\}}{\#\{t \in \mathcal{Y} : |t| = n\}}$$

$$\mu^+\left(\frac{\mathcal{X}}{\mathcal{Y}}\right) = \limsup_{n \to \infty} \frac{\#\{t \in \mathcal{X} : |t| = n\}}{\#\{t \in \mathcal{Y} : |t| = n\}}$$

For technical reasons, we assume $\frac{0}{0} := 1$. Given any subclass $\mathcal{X} \subseteq \mathcal{C}$ and $n \in \mathbb{N}$, we denote by $\mathcal{X}_n$ the set of all combinators from $\mathcal{X}$ that are of size n. Obviously $\mathcal{X}_n$ is always finite.

4   Generating Functions

Many questions concerning the asymptotic behavior of sequences of real non-negative numbers can be efficiently resolved by analyzing the behavior of their generating functions (see [18] for introductory reference). This is the approach we take to determine the asymptotic fraction of certain combinatory logic terms. Let $(a_n)_{n \in \mathbb{N}}$ be a sequence of non-negative numbers. The power series $A(z) = \sum_{n \in \mathbb{N}} a_n z^n$ is called the *generating function* enumerating the sequence $(a_n)_{n \in \mathbb{N}}$. We denote by $[z^n]\{A(z)\}$ the coefficient of z^n in the expansion of $A(z)$. We say that two sequences $(A_n)_{n \in \mathbb{N}}$ and $(B_n)_{n \in \mathbb{N}}$ are *asymptotically equivalent* if $\lim_{n \to \infty} \frac{A_n}{B_n} = 1$. In such a case we write $A_n \sim B_n$. The following theorem is a well-known result in the theory of generating functions. Its derivation from Szegö Lemma (see [16]) can be found, e.g., in [21, Theorem 22].

Theorem 1 (Generating Function Method via Szegö Lemma). *Let A, B be functions satisfying the following conditions:*

1. *A, B are analytic in $|z| < 1$ with $z = 1$ being the only singularity on the circle $|z| = 1$,*
2. *A, B have the following expansions in the vicinity of $z = 1$:*

$$A(z) = \sum_{p \geq 0} a_p (1 - z)^{\frac{p}{2}}, \qquad B(z) = \sum_{p \geq 0} b_p (1 - z)^{\frac{p}{2}},$$

   where $b_1 \neq 0$.

Let $\tilde{A}$ and $\tilde{B}$ be functions satisfying $\tilde{A}(\sqrt{1 - z}) = A(z)$ and $\tilde{B}(\sqrt{1 - z}) = B(z)$. Then

$$\lim_{n \to \infty} \frac{[z^n]\{A(z)\}}{[z^n]\{B(z)\}} = \frac{a_1}{b_1} = \frac{\tilde{A}'(0)}{\tilde{B}'(0)}.$$

Theorem 2 (Pringsheim, see [5, Theorem IV.6]). *If $A(z)$ is representable at the origin by a series expansion that has non-negative coefficients and radius of convergence R, then the point $z = R$ is a singularity of $A(z)$.*

Theorem 3 (Exponential Growth Formula, see [5, Theorem IV.7]). *If $A(z)$ is analytic at 0 and R is the modulus of a singularity nearest to the origin in the sense that*

$$R = \sup\{r \geq 0 \; : \; A \text{ is analytic in } |z| < r\},$$

then the coefficient $a_n = [z^n]\{A(z)\}$ satisfies

$$a_n = R^{-n}\theta(n) \quad with \quad \limsup |\theta(n)|^{\frac{1}{n}} = 1.$$

By T_n we denote n-th Catalan number, i.e., the number of expressions (or equivalently trees) containing n pairs of parentheses which are correctly matched. It is well-known that $T_n = \frac{1}{n+1}\binom{2n}{n}$ and that

$$\lim_{n \to \infty} \frac{T_{n+1}}{T_n} = 4. \tag{1}$$

Since the set of all combinators is defined by a very simple grammar, we can easily count all combinators of a given size.

Fact 1 (see, e.g. [4]). *Let C be the generating function enumerating combinators of a given size. Then*

$$C(z) = \frac{1 - \sqrt{1 - 8z}}{2z} \quad and \quad |\mathcal{C}_n| = [z^n]\{C(z)\} = 2^{n+1} \cdot T_n$$

5 Weakly Normalizing Combinators

Let us start with the well-known classical fact observed already in [3].

Theorem 4 (Standarization Theorem). *If a combinator is weakly normalizing, then the leftmost outermost reduction always leads to a normal form.*

Another classical observation is that the set of weakly normalizing combinators $\mathcal{WN}$ is undecidable. It follows that there is no purely syntactic formula enumerating $\mathcal{WN}_n$ and thus we cannot find the cardinality of $\mathcal{WN}_n$ explicitly. For that reason we take the following approach. We find feasible subclasses of $\mathcal{WN}$ and $\mathcal{C} \setminus \mathcal{WN}$ and use them to bound the density of $\mathcal{WN}$ in $\mathcal{C}$.

Lemma 1. *Asymptotically at least $\frac{1}{32}$ of combinators are weakly normalizing i.e., $\mu^- \left(\frac{\mathcal{WN}}{\mathcal{C}} \right) \geq \frac{1}{32}$.*

Proof. Let $\mathcal{L}$ be the class of combinators which are either of the form $\mathbf{KK}M$ or $\mathbf{KS}M$, where $M \in \mathcal{C}$ is an arbitrary combinator. Let us notice that in just one reduction step every combinator from this class is reducible either to $\mathbf{K}$ or to $\mathbf{S}$. Therefore $\mathcal{L}$ is a subset of all weakly normalizing combinators. Moreover, we have $|\mathcal{L}_n| = 2|\mathcal{C}_{n-2}|$ for $n \geq 2$ and so

$$\mu \left(\frac{\mathcal{L}}{\mathcal{C}} \right) = \lim_{n \to \infty} \frac{|\mathcal{L}_n|}{|\mathcal{C}_n|} = \lim_{n \to \infty} \frac{2|\mathcal{C}_{n-2}|}{|\mathcal{C}_n|} = \lim_{n \to \infty} \frac{2^n \cdot T_{n-2}}{2^{n+1} \cdot T_n} \overset{(1)}{=} \frac{1}{2} \cdot \frac{1}{4^2} = \frac{1}{32}. \qquad \square$$

Lemma 2. *Asymptotically at most $1 - \frac{1}{2^{18}}$ of combinators are weakly normalizing, i.e., $\mu^+ \left(\frac{\mathcal{WN}}{\mathcal{C}} \right) \leq 1 - \frac{1}{2^{18}}$.*

Proof. Let $\omega_1 = \mathbf{S(SS)SSSS}$ and $\omega_2 = \mathbf{SSS(SS)SS}$. Consider the class $\mathcal{U}$ of combinators that are in form of $\omega_1 M_1 \ldots M_k$ or $\omega_2 M_1 \ldots M_k$ for arbitrary $k \geq 0$ and $M_1, \ldots, M_k \in \mathcal{C}$. In [1] it was shown that ω_1 is not normalizable, which implies that so is ω_2 since ω_1 reduces to ω_2. By the standarization theorem for combinatory logic, we obtain that $\mathcal{U} \subseteq (\mathcal{C} \setminus \mathcal{WN})$. Since both ω_1 and ω_2 are of size 6, we get $|\mathcal{U}_n| = 2^{n-6} \cdot T_{n-6}$ for $n \geq 7$. Finally,

$$\mu \left(\frac{\mathcal{U}}{\mathcal{C}} \right) = \lim_{n \to \infty} \frac{|\mathcal{U}_n|}{|\mathcal{C}_n|} = \lim_{n \to \infty} \frac{|\mathcal{C}_{n-6}|}{|\mathcal{C}_n|} = \lim_{n \to \infty} \frac{T_{n-6}}{2^6 \cdot T_n} \overset{(1)}{=} \frac{1}{2^6} \cdot \frac{1}{4^6} = \frac{1}{2^{18}}. \qquad \square$$

The two above lemmas show that the density of weakly normalizing combinators, provided it exists, is neither zero nor one. In [4] it was shown that a random combinator is not strongly normalizing. This immediately implies the following result.

Theorem 5. *Asymptotically almost all weakly normalizing terms are not strongly normalizing i.e., $\mu \left(\frac{\mathcal{SN}}{\mathcal{WN}} \right) = 0$.*

6 Combinatorial Results

Lemma 3. *Let $\mathcal{N}$ be the set of combinators in normal form. The generating function $F_{\mathcal{N}}$ enumerating cardinality of $\mathcal{N}$ is given by*

$$F_{\mathcal{N}}(z) = \frac{1 - 2z - \sqrt{1 - 4z - 4z^2}}{2z^2}.$$

This implies

$$[z^n]\{F_{\mathcal{N}}(z)\} \sim (2 + 2\sqrt{2})^n \theta(n) = (4.82843\ldots)^n \theta(n)$$

with $\limsup_{n\to\infty} |\theta(n)|^{\frac{1}{n}} = 1$.

Proof. The grammar for $\mathcal{N}$ is given by

$$\mathcal{N} := \mathbf{S} \mid \mathbf{K} \mid \mathbf{K}\,\mathcal{N} \mid \mathbf{S}\,\mathcal{N} \mid \mathbf{S}\,\mathcal{N}\,\mathcal{N}.$$

It follows that the generating function $F_{\mathcal{N}}$ satisfies

$$F_{\mathcal{N}}(z) = 2 + 2zF_{\mathcal{N}}(z) + z^2(F_{\mathcal{N}}(z))^2.$$

Solving the equation for $F_{\mathcal{N}}(z)$ we obtain two solutions $\frac{1-2z\pm\sqrt{1-4z-4z^2}}{2z^2}$. Because $\lim_{n\to\infty} F_{\mathcal{N}}(0) = 2$ we conclude that $F_{\mathcal{N}}(z) = \frac{1-2z-\sqrt{1-4z-4z^2}}{2z^2}$. In order to compute the asymptotic growth of $[z^n]\{F_{\mathcal{N}}(z)\}$, we start with the observation that $F_{\mathcal{N}}(z)$ has an analytic continuation in 0 and its radius of convergence R is equal to $\frac{1}{2}(\sqrt{2} - 1)$. By Pringsheim's theorem be obtain that R is also the modulus of the dominating singularity of $F_{\mathcal{N}}(z)$ and thus applying the Exponential Growth Formula we obtain that

$$[z^n]\{F_{\mathcal{N}}(z)\} = R^{-n}\theta(n) \sim (4.82843\ldots)^n\theta(n) \text{ with } \limsup_{n\to\infty} |\theta(n)|^{\frac{1}{n}} = 1. \qquad \square$$

In order to determine the density of normal forms in the set of all strongly normalizing combinators, we define a class $\mathcal{G}$ as the set of combinators defined by the following grammar:

$$\mathcal{G} := \mathbf{S} \mid \mathbf{K} \mid \mathbf{KKK} \mid \mathbf{K}\,\mathcal{G} \mid \mathbf{S}\,\mathcal{G} \mid \mathbf{S}\,\mathcal{G}\,\mathcal{G}.$$

Since $\mathcal{G}$ contains all productions of $\mathcal{N}$, we have $\mathcal{N} \subseteq \mathcal{G}$. Moreover, the only redexes in $\mathcal{G}$ are of the form $\mathbf{KKK}$, which implies that $\mathcal{G} \subseteq \mathcal{SN}$.

Lemma 4. *The generating function $F_{\mathcal{G}}$ enumerating cardinality of $\mathcal{G}$ is given by*

$$F_{\mathcal{G}}(z) = \frac{1 - 2z - \sqrt{1 - 4z - 4z^2 - 4z^4}}{2z^2},$$

which yields

$$[z^n]\{F_{\mathcal{G}}(z)\} \sim (4.85823\ldots)^n \theta(n)\, with \limsup |\theta(n)|^{\frac{1}{n}} = 1.$$

Proof. Given the grammar for $\mathcal{G}$ we obtain that $F_{\mathcal{G}}$ satisfies the following equation $F_{\mathcal{G}}(z) = 2 + z^2 + 2zF_{\mathcal{G}}(z) + z^2(F_{\mathcal{G}}(z))^2$. Solving for $F_{\mathcal{G}}(z)$ we find two possible solutions $\frac{1-2z\pm\sqrt{1-4z-4z^2-4z^4}}{2z^2}$. Since $\lim_{n\to\infty} F_{\mathcal{G}}(0) = 2$ we conclude that $F_{\mathcal{G}}(z) = \frac{1-2z-\sqrt{1-4z-4z^2-4z^4}}{2z^2}$. In order to find the dominating singularity of $F_{\mathcal{G}}(z)$, we examine the real roots of $\sqrt{1-4z-4z^2-4z^4}$. This expression yields two real roots $z_1 \approx -0.800151$ and $z_2 \approx 0.205836$. Because z_2 lies closer to the origin, it follows that $z \approx 0.205836$ dictates the asymptotic growth of $[z^n]\{F_{\mathcal{G}}(z)\}$. Applying the Exponential Growth Formula we obtain that

$$[z^n]\{F_{\mathcal{G}}(z)\} \sim (4.85823\ldots)^n\theta(n) \text{ with } \limsup|\theta(n)|^{\frac{1}{n}} = 1. \qquad \square$$

Theorem 6. *Asymptotically almost all strongly normalizing terms are not in normal form, i.e., $\mu\left(\frac{\mathcal{N}}{\mathcal{SN}}\right) = 0$.*

Proof. Similarly to $\mathcal{WN}$, the set of strongly normalizing combinators $\mathcal{SN}$ is undecidable and therefore we cannot enumerate $\mathcal{SN}_n$ explicitly. Fortunately, it suffices to prove that $\mu\left(\frac{\mathcal{N}}{\mathcal{H}}\right) = 0$ for some sufficiently large subclass $\mathcal{H} \subseteq \mathcal{SN}$. Let $\mathcal{G}$ be the set of combinators as defined in Lemma 4. We claim that $\mu\left(\frac{\mathcal{N}}{\mathcal{G}}\right) = 0$. Indeed, using Lemmas 3 and 4 we obtain

$$\mu\left(\frac{\mathcal{N}}{\mathcal{G}}\right) = \lim_{n\to\infty}\frac{[z^n]\{F_{\mathcal{N}}(z)\}}{[z^n]\{F_{\mathcal{G}}(z)\}} = \lim_{n\to\infty}\frac{(4.82843\ldots)^n}{(4.85823\ldots)^n} = 0. \qquad \square$$

Lemma 5. *Let $t_0 \in \mathcal{N}$ be a combinator of size $|t_0| \geq 1$. Let $\mathcal{N}_{t_0}$ be the set of combinators in normal form which contain t_0 as a subterm. The generating function $F_{\mathcal{N}_{t_0}}$ enumerating cardinality of $\mathcal{N}_{t_0}$ is given by*

$$F_{\mathcal{N}_{t_0}}(z) = \frac{-\sqrt{1-4z-4z^2}+\sqrt{1-4z-4z^2+4z^{|t_0|+2}}}{2z^2}.$$

Proof. Note that if $Q \in \mathcal{N}_{t_0}$ then either:

1. $Q = t_0$, or
2. $Q = \mathbf{K}M$ and t_0 is a subterm of M, or
3. $Q = \mathbf{S}M$ and t_0 is a subterm of M, or
4. $Q = \mathbf{S}MP$ and t_0 is a subterm of M but not P, or
5. $Q = \mathbf{S}MP$ and t_0 is a subterm of P but not M, or
6. $Q = \mathbf{S}MP$ and t_0 is a subterm of both M and P.

It follows that $F_{\mathcal{N}_{t_0}}$ satisfies the following equation

$$F_{\mathcal{N}_{t_0}}(z) = z^{|t_0|} + 2zF_{\mathcal{N}_{t_0}}(z) + 2z^2(F_{\mathcal{N}}(z) - F_{\mathcal{N}_{t_0}}(z))F_{\mathcal{N}_{t_0}}(z) + z^2(F_{\mathcal{N}_{t_0}}(z))^2.$$

Using the generating function for $\mathcal{N}$ we solve this equation for $F_{\mathcal{N}_{t_0}}(z)$ and obtain two solutions

$$\frac{-\sqrt{1-4z-4z^2}\pm\sqrt{1-4z-4z^2+4z^{|t_0|+2}}}{2z^2}.$$

Since there is no $Q \in \mathcal{N}_{t_0}$ of size 0, we get that $\lim_{z \to 0} F_{\mathcal{N}_{t_0}}(z) = 0$ and finally

$$F_{\mathcal{N}_{t_0}}(z) = \frac{-\sqrt{1 - 4z - 4z^2} + \sqrt{1 - 4z - 4z^2 + 4z^{|t_0|+2}}}{2z^2}.$$

$\square$

Theorem 7. *Let $t_0 \in \mathcal{N}$. The density of combinators in normal form which contain t_0 as a subterm is 1.*

Proof. We prove this result applying Theorem 1. We start with normalizing $F_{\mathcal{N}_{t_0}}$ and $F_{\mathcal{N}}$ in such a way that both generating functions are analytic in the disc $|z| < 1$ with $z = 1$ being their only singularity on the circle $|z| = 1$. For convenience, let us shift both generating functions by two positions obtaining $\hat{F}_{\mathcal{N}}(z) := z^2 F_{\mathcal{N}}(z)$ and $\hat{F}_{\mathcal{N}_{t_0}}(z) := z^2 F_{\mathcal{N}_{t_0}}(z)$. Since $R = \frac{1}{2}(\sqrt{2} - 1)$ is the dominating singularity of both $F_{\mathcal{N}}$ and $F_{\mathcal{N}_{t_0}}$, we define $\overline{F}_{\mathcal{N}}(z) := \hat{F}_{\mathcal{N}}(Rz)$ and $\overline{F}_{\mathcal{N}_{t_0}}(z) := \hat{F}_{\mathcal{N}_{t_0}}(Rz)$.

First, we examine $\overline{F}_{\mathcal{N}}(z)$. Simplifying, we get

$$\overline{F}_{\mathcal{N}}(z) = \frac{1}{2}\left(-\sqrt{2}z + z - \sqrt{(1 - z)\left(1 - (2\sqrt{2} - 3)z\right)} + 1\right).$$

Note that the inner expression $\sqrt{(1 - z)\left(1 - (2\sqrt{2} - 3)z\right)}$, carrying the singularities of $\overline{F}_{\mathcal{N}}(z)$, has exactly two roots, i.e., $z_1 = 1$ and $z_2 = \frac{1}{2\sqrt{2}-3} \approx -5.82843$. It follows that $\overline{F}_{\mathcal{N}}(z)$ must be analytic in the disc $|z| < 1$ with $z = 1$ being the only singularity on the circle $|z| = 1$. Moreover, $\overline{F}_{\mathcal{N}}(z)$ yields an expansion in the vicinity of $z = 1$ in form of $\sum_{p \geq 0} w_p (1 - z)^{p/2}$ with $w_1 = -\frac{1}{2}\sqrt{4 - 2\sqrt{2}} \neq 0$.

Simplifying $\overline{F}_{\mathcal{N}_{t_0}}$, we obtain

$$\overline{F}_{\mathcal{N}_{t_0}} = -\frac{1}{2}\sqrt{(1 - z)\left(1 - (2\sqrt{2} - 3)z\right)}$$
$$+ \frac{1}{2}\sqrt{1 - 2(\sqrt{2} - 1)z - (3 - 2\sqrt{2})z^2 + 2^{-|t_0|}(\sqrt{2} - 1)^{|t_0|+2}}.$$

Since $1 - 2(\sqrt{2} - 1)z - (3 - 2\sqrt{2})z^2 + 2^{-|t_0|}(\sqrt{2} - 1)^{|t_0|+2}$ is decreasing in $[0, 1]$ attaining values 1 and 0 for $z = 0$ and $z = 1$ respectively, we obtain that $\overline{F}_{\mathcal{N}_{t_0}}$ is analytic in the disc $|z| < 1$ with $z = 1$ being the only singularity on the circle $|z| = 1$. Moreover, $\overline{F}_{\mathcal{N}_{t_0}}$ has an expansion in the vicinity of $z = 1$ in form of $\sum_{p \geq 0} v_p (1 - z)^{p/2}$ with $v_1 = -\frac{1}{2}\sqrt{4 - 2\sqrt{2}}$.

Next, let us consider functions $\widetilde{F}_{\mathcal{N}}$ and $\widetilde{F}_{\mathcal{N}_{t_0}}$ such that

$$\widetilde{F}_{\mathcal{N}}(\sqrt{1 - z}) = \overline{F}_{\mathcal{N}}(z) \quad \text{and} \quad \widetilde{F}_{\mathcal{N}_{t_0}}(\sqrt{1 - z}) = \overline{F}_{\mathcal{N}_{t_0}}(z).$$

By the analyticity of $\overline{F}_{\mathcal{N}}$ and $\overline{F}_{\mathcal{N}_{t_0}}$ in the disc $|z| < 1$, we obtain that both $\widetilde{F}_{\mathcal{N}}(z)$ and $\widetilde{F}_{\mathcal{N}_{t_0}}(z)$ yield derivatives in this disc and thus $\widetilde{F}'_{\mathcal{N}}(0)$ and $\widetilde{F}'_{\mathcal{N}_{t_0}}(0)$ exist.

Finally, computing those derivatives we get $\widetilde{F}'_{\mathcal{N}}(0) = \widetilde{F}'_{\mathcal{N}_{t_0}}(0) = -\frac{1}{2}\sqrt{4 - 2\sqrt{2}}$ and so

$$\lim_{n \to \infty} \frac{[z^n]\{F_{\mathcal{N}_{t_0}}\}}{[z^n]\{F_{\mathcal{N}}\}} = \lim_{n \to \infty} \frac{R^{-n-2}[z^{n+2}]\{\overline{F}_{\mathcal{N}_{t_0}}\}}{R^{-n-2}[z^{n+2}]\{\overline{F}_{\mathcal{N}}\}} = \lim_{n \to \infty} \frac{\widetilde{F}'_{\mathcal{N}_{t_0}}(0)}{\widetilde{F}'_{\mathcal{N}}(0)} = 1. \qquad \square$$

Theorem 8. *Asymptotically almost all normal forms are not typeable, i.e.,* $\mu\left(\frac{\mathcal{TYP} \cap \mathcal{N}}{\mathcal{N}}\right) = 0.$

Proof. Note that $\Omega = \mathbf{S}(\mathbf{SKK})(\mathbf{SKK})$ is in normal form and is not typeable. Directly from Theorem 7 we obtain that asymptotically almost every combinator in normal form contains Ω as a subterm and is thus not typeable. $\qquad \square$

References

1. Barendregt, H.P., Bergstra, J., Klop, J.W., Volken, H.: Some notes on lambda reduction, in: Degrees, reductions and representability in the lambda calculus. Preprint no. 22, University of Utrecht, Department of mathematics, pp. 13–53 (1976)
2. Chauvin, B., Flajolet, P., Gardy, D., Gittenberger, B.: And/Or trees revisited. Comb. Probab. Comput. **13**(4–5), 475–497 (2004)
3. Curry, H., Feys, R.: Combinatory Logic, vol. I. North Holland, Amsterdam (1958)
4. David, R., Grygiel, K., Kozik, J., Raffalli, C., Theyssier, G., Zaionc, M.: Asymptotically almost all λ-terms are strongly normalizing. Logical Methods Comput. Sci. **9**, 1–30 (2013)
5. Flajolet, P., Sedgewick, R.: Analytic Combinatorics. Cambridge University Press, Cambridge (2009)
6. Fournier, H., Gardy, D., Genitrini, A., Zaionc, M.: Classical and intuitionistic logic are asymptotically identical. In: Duparc, J., Henzinger, T.A. (eds.) CSL 2007. LNCS, vol. 4646, pp. 177–193. Springer, Heidelberg (2007)
7. Gardy, D. Random boolean expressions. In: Discrete Mathematics and Theoretical Computer Science Proceedings AF, pp.1–36 (2005)
8. Gardy, D., Woods, A.: And/Or tree probabilities of boolean functions. Discrete Math. Theor. Comput. Sci. **6**, 139–146 (2005)
9. Genitrini, A., Kozik, J.: In the full propositional logic, 5/8 of classical tautologies are intuitionistically valid. Ann. Pure Appl. Logic **163**(7), 875–887 (2012)
10. Grygiel, K., and Lescanne, P. Counting terms in the binary lambda calculus. In: DMTCS 25th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (2014)
11. Kostrzycka, Z., Zaionc, M.: Statistics of intuitionistic versus classical logic. Stud. Logica **76**(3), 307–328 (2004)
12. Kozik, J.: Subcritical pattern languages for And/Or trees. In: DMTCS Proceedings from Fifth Colloquium on Mathematics and Computer Science Algorithms Trees, Combinatorics and Probabilities, pp. 437–448 (2008)
13. Lefmann, H., Savický, P.: Some typical properties of large And/Or Boolean formulas. Random Struct. Algorithms **10**, 337–351 (1997)
14. Moczurad, M., Tyszkiewicz, J., Zaionc, M.: Statistical properties of simple types. Math. Struct. Comput. Sci. **10**(5), 575–594 (2000)

15. Olivier, B., Danièle, G., Bernhard, G., Alice, J.: Enumeration of generalized BCI lambda-terms. Electr. J. Comb. **20**, 4 (2013)
16. Szegö, G.: Orthogonal polynomials. Am. Math. Soc. Colloquium Ser. Publ. **23**, 413–421 (1967)
17. Tromp, J. Binary lambda calculus and combinatory logic. Unpublished manuscript (2014). http://tromp.github.io/cl/LC.pdf.
18. Wilf, H.: Generating Functionology. Academic Press, Boston (1994)
19. Woods, A.: On the probability of absolute truth for And/Or formulas. Bull. Symbolic Logic **12**, 3 (2006)
20. Zaionc, M.: On the asymptotic density of tautologies in logic of implication and negation. Rep. Math. Logic **39**, 67–87 (2005)
21. Zaionc, M.: Probability distribution for simple tautologies. Theor. Comput. Sci. **355**(2), 243–260 (2006)