Under consideration for publication in Math. Struct. in Comp. Science

Statistical Properties of Simple Types

M. MOCZURAD,[1] J. TYSZKIEWICZ[2†] and M. ZAIONC[1‡]

[1] *Computer Science Department, Jagiellonian University,*
Nawojki 11, 30-072 Krakow, Poland.
E-mail madry@softlab.ii.uj.edu.pl, zaionc@ii.uj.edu.pl
[2] *Institute of Informatics, University of Warsaw, Banacha 2, 02-097 Warszawa, Poland,*
E-mail jty@mimuw.edu.pl *on leave at School CSE, UNSW, 2052 Sydney, Australia.*

Received April 1999

This work is in honour of Roger Hindley on his $60^{\mathbf{th}}$ birthday.

We consider types and typed lambda calculus over a finite number of ground types. We are going to investigate the size of the fraction of inhabited types of the given length n against the number of all types of length n. The plan of this paper is to find the limit of that fraction when $n \to \infty$. The answer to this question is equivalent to finding the "density" of inhabited types in the set of all types, or the so-called asymptotic probability of finding an inhabited type in the set of all types. Under the Curry-Howard isomorphism this means finding the density or asymptotic probability of provable intuitionistic propositional formulas in the set of all formulas. For types with one ground type (formulas with one propositional variable) we prove that the limit exists and is equal to $1/2 + \sqrt{5}/10$, which is approximately 72.36%. This means that a long random type (formula) has this probability to be inhabited (tautology). We also prove that for every finite number k of ground-type variables, the density of inhabited types is always positive and lies between $(4k+1)/(2k+1)^2$ and $(3k+1)/(k+1)^2$. Therefore we can easily see that the density is decreasing to 0 with k going to infinity. From the lower and upper bounds presented we can deduce that at least $1/3$ of classical tautologies are intuitionistic.

† This author has been supported by the Australian Research Council grant A 49800112 (1998–2000) and by the Polish KBN, grant 8T11C02716.
‡ The first and the third author have been partially supported by the Polish KBN, grant 8T11C01816.

1. Introduction

In this paper we examine the density of inhabited types among all first order types of a given length. A type is inhabited, if there is at least one closed term of λ-calculus of this type. The inhabitation problem for a simple typed λ-calculus, which became the real motivation for this paper, was deeply studied by Hindley and is fully analyzed in (Hindley 1997). By the Curry-Howard isomorphism, inhabitation is equivalent to the statement that the type is an intuitionistic tautology (in the language of pure implication). Thus what we indeed attempt to measure is the density of intuitionistic tautologies among all purely implicative formulas. In this statement the λ-calculus is not even mentioned. What we however borrow from it is our interest in formulas with very few propositional variables, which corresponds to the common situation with just a few basic types. Indeed, λ-calculus is a (or perhaps even *the*) mathematical model for functional programming, and in programming languages the number of atomic types is usually pretty small. In the logical setting such restriction would be perhaps questionable, but given our interest in λ-calculus, it is quite natural and legitimate.

Our first main result is a precise estimate of the number of intuitionistic tautologies among formulas of length n with one propositional variable. This result was partially motivated by a short note at the end of Statman's paper (Statman 1980), in which he estimated the fraction of inhabited types as lying between 0.625 and 0.86, and ended his notice with the nice sentence: *"It is a good bet but not a sure thing, that ρ (type) contains a closed term."* As we prove, this is a really *good bet* because among types with one variable the fraction of tautologies tends to $1/2 + \sqrt{5}/10$ as the length of types tends to infinity.

For types with $k > 1$ atomic types we do not know the exact fraction of tautologies, but we prove two bounds which almost match. After isolating special formulas, which we call simple tautologies and simple nontautologies, we prove that their fractions among all formulas of length n approach $(4k+1)/(2k+1)^2$ and $k(k-1)/(k+1)^2$, respectively. So, for large k, the upper bound on the fraction of tautologies is about 3 times the lower bound. We conjecture that indeed the fraction of tautologies, for large k, is very close to the lower bound. In particular, we conjecture that most of the tautologies are indeed simple (in our terminology).

Since our simple tautologies are classical tautologies, and simple nontautologies are classically invalid, our conjecture implies that for large k almost all classical tautologies are intuitionistic, and in fact simple. At present we know only that at least one third of classical tautologies are intuitionistic ones (see Theorem 8.1).

Finally, it is worth observing that, in the light of our results, the number of tautologies with k variables eventually decreases with k: if k grows more than by a factor of 3, the (asymptotic) fraction of tautologies must drop below the old value. We prove separately that asymptotically there are indeed fewer tautologies with 2 atomic types than with 1 atomic type.

2. Prerequisites

2.1. λ-*calculus with simple types*

We shall consider a simple typed λ-calculus with a finite number of ground types. Types are defined as follows: every ground type is a type and if τ and μ are types then $(\tau \to \mu)$ is a type. We will use the following notation: if $\tau_1,\ \tau_2,\ \ldots,\ \tau_n,\ \mu$ are types then by $\tau_1, \tau_2, \ldots, \tau_n \to \mu$ we mean the type $\tau_1 \to (\tau_2 \to (\ldots \to (\tau_n \to \mu)\ldots))$. Therefore, every type τ has the form $\tau_1, \tau_2, \ldots, \tau_n \to \alpha$ where α is a ground type. If τ has the form $\tau_1, \ldots, \tau_n \to \alpha$ then τ_i for $i \leq n$ is called the i-th component/premise of the type τ and is denoted by $\tau[i]$. This notation can be iterated so by a type $\tau[i_1, \ldots, i_{k-1}, i_k]$ we mean the type $(\tau[i_1, \ldots, i_{k-1}])[i_k]$. By the length of a type τ we mean the number of occurrences of ground types in τ.

For any type τ a denumerable set $V(\tau)$ of variables of this type is given. Any variable of type τ is a term of type τ. If T is a term of type $\tau \to \mu$ and S is a term of type τ, then TS is a term of type μ. If T is a type μ term and x is a type τ variable, then $\lambda x.T$ is a term of type $\tau \to \mu$. The axioms of equality between terms have the form of $\beta\eta$ conversions and the convertible terms will be written as $T =_{\beta\eta} S$. Term T is in long normal form if $T = \lambda x_1 \ldots x_n.yT_1 \ldots T_k$, where y is an x_i for some $i \leq n$ or y is a free variable, T_j for $j \leq k$ is in long normal form and $yT_1\ \ldots\ T_k$ is a ground type term. Long normal forms exist and are unique up to α conversion (see (Statman 1980, p. 533) or (Hindley 1997, p. 110)). A closed term is a term without free variables.

Definition 2.1. A type is called *inhabited* if there is a closed term of this type.

For very detailed analysis of inhabited types see (Hindley 1997).

The following Lemma belongs to the folklore:

Lemma 2.2. Any intuitionistic tautology is a classical tautology.

2.2. *Generating functions*

The main tool we use for dealing with asymptotics of sequences of numbers are *generating functions.* A nice exposition of the method can be found in (Wilf 1994; Comtet 1974).

Our main task in this paper is to determine limits of various sequences of real numbers. For this purpose combinatorics has developed an extremely powerful tool, in the form of generating series and generating functions.

Let $A = A_0, A_1, A_2, \ldots$ be a sequence of real numbers (if it is necessary, we tacitly extend A to the left, by assuming $A_n = 0$ for negative n). The *ordinary generating series* for A is the formal power series $\sum_{n=0}^{\infty} A_n z^n$. And, of course, formal power series are in one-to-one correspondence to sequences. However, considering z as a complex variable, this series, as known from the theory of analytic functions, converges uniformly to a function $f_A(z)$ in some open disc $\{z \in \mathrm{C} : |z| < R\}$ of maximal diameter, and $R \geq 0$ is called its radius of convergence. So with the sequence A we can associate a complex function $f_A(z)$, called the *ordinary generating function* for A, defined in a neighbourhood of 0. This correspondence is one-to-one again (unless $R = 0$), since, as it is well known from the theory of analytic functions, the expansion of a complex function $f(z)$, analytic in a neighbourhood of z_0, into a power series $\sum_{n=0}^{\infty} A_n(z - z_0)^n$ is unique, and, moreover, this series is the Taylor series, given by

$$A_n = \frac{1}{n!}\frac{d^n f}{dz^n}(z_0). \tag{1}$$

Many questions concerning the asymptotic behaviour of A can be efficiently resolved by analysing the behaviour of f_A at the complex circle $|z| = R$.

This is the approach we take to determine the asymptotic fraction of intuitionistic tautologies among all types of a given length.

The key tool will be the following result, in which the symbol $[z^n]\{F\}$ stands for the coefficient of z^n in the exponential series expansion of F.) It is due to Szegö (Szegö 1975, Thm. 8.4), see as well (Wilf 1994, Thm. 5.3.2).

Theorem 2.3 (Szegö lemma). Let $v(z)$ be analytic in $|z| < 1$ with a finite number of singularities $e^{i\varphi^{(k)}}$, $k = 1, \ldots, s$, at the circle $|z| = 1$. Suppose that in the neighbourhood of each $e^{i\varphi^{(k)}}$, $v(z)$ has the expansion of the form

$$v(z) = \sum_{p \geq 0} v_p^{(k)} (1 - z e^{-i\varphi^{(k)}})^{a^{(k)} + p b^{(k)}}, \tag{2}$$

where $a^{(k)} \in \mathrm{C}$ and $b^{(k)} > 0$ is real, and the branch chosen above for the expansion equals $v(0)$ for $z = 0$. Then

$$[z^n]\{v(z)\} = \sum_{k=1}^{s} \sum_{p=0}^{\xi(q)} v_p^{(k)} \binom{a^{(k)} + p b^{(k)}}{n} (-e^{i\varphi^{(k)}})^n + O(n^{-q}), \tag{3}$$

with

$$\xi(q) = \max_{k=1\ldots s} \lceil (1/b^{(k)})(q - \Re(a^{(k)}) - 1) \rceil. \tag{4}$$

In (3) $\binom{a}{n}$ stands for $a(a-1)\ldots(a-(n-1))/n!$.

As the reader may now expect, while working with propositional intuitionistic logic and simply typed λ-calculus, we will be often concerned with complex analysis, analytic functions and their singularities.

In all our applications we will indeed have only one singularity $z = 1$ with $b = 1/2$, and we will be satisfied with error bound $O(n^{-2})$. It follows that our ξ is 2. Moreover, $\binom{0}{n} = \binom{1}{n} = 0$ for $n > 1$. Therefore we will be using much simpler version of Szegö Lemma:

Corollary 2.4. Let $v(z)$ be analytic in $|z| < 1$ with $z = 1$ the only singularity at the circle $|z| = 1$. If $v(z)$ in the vicinity of $z = 1$ has the expansion of the form

$$v(z) = \sum_{p \geq 0} v_p (1 - z)^{p/2}, \tag{5}$$

where $b > 0$, and the branch chosen above for the expansion equals $v(0)$ for $z = 0$, then

$$[z^n]\{v(z)\} = v_1 \binom{1/2}{n} (-1)^n + O(n^{-2}). \tag{6}$$

3. Counting types

In this section we present some properties of the Catalan numbers characterising the number of possible bracketings of a type and consequently, the number of types of a given length.

Let us recall the definition of the Catalan number.

Definition 3.1. The Catalan number C_n is given by the recursion:

$$C_0 = 0 \qquad C_1 = 1 \qquad C_n = \sum_{i=1}^{n-1} C_i C_{n-i}. \tag{7}$$

Lemma 3.2. The number of types of length n with one ground type is C_n.

Proof. A type of length n with one ground type α is a bracketing of the sequence $\underbrace{\alpha \to \alpha \to \ldots \to \alpha}_{n}$. The number of such bracketings is immediately seen to satisfy the recursion (7). ☐

For more elaborate treatment of Catalan numbers see (Wilf 1994, pp. 43–44). We mention only the following is a well-known nonrecursive formula for C_n.

$$C_n = \frac{1}{n}\binom{2n-2}{n-1}, \tag{8}$$

and quote some properties which are its consequences.

For every $n \geq 1$ and for every $k \geq 1$

$$\frac{C_n}{C_{n+1}} = \frac{1}{4} + \frac{3}{8n-4}, \tag{9}$$

$$\frac{C_n}{C_{n+k}} > \frac{1}{4^k}, \tag{10}$$

$$\lim_{n\to\infty} \frac{C_n}{C_{n+k}} = \frac{1}{4^k}. \tag{11}$$

Definition 3.3. By $C_n(p)$ we mean the number of bracketings of a type of length n having p premises, i.e. formulas which are of the form: $\tau = \tau_1 \to (\ldots \to (\tau_p \to \alpha))$, where α is a ground type. Since numbers $C_n(p)$ are the cardinalities of disjoint sets of formulas for different p's and since there are no types of length n having more then $n-1$ premises, for $n \geq 2$ we have:

$$C_n = C_n(1) + \cdots + C_n(n-1).$$

Lemma 3.4. Number $C_n(p)$ is given by the recursion (on p):

$$\begin{aligned} C_n(1) &= C_{n-1}, \\ C_n(p) &= \sum_{i=1}^{n-p} C_i C_{n-i}(p-1). \end{aligned}$$

Proof. It is obvious that $C_n(1) = C_{n-1}$, since $C_n(1)$ is the number of types of the form $\tau \to \alpha$. Consider

$$\tau = \tau_1 \to \underbrace{(\tau_2 \to (\ldots(\tau_p \to \alpha)\ldots))}_{\mu},$$

where τ_1 is of length i. The number of possible bracketings of τ is the number of bracketings of τ_1 (i.e. C_i) and μ (i.e. $C_{n-i}(p-1)$), summed over all possible divisions at position i. The summation stops at $i = n-p$, since beginning with $i = n-p+1$ the terms become zero. $\square$

Lemma 3.5. For all $p > 0$ holds $\lim_{n\to\infty} C_n(p)/C_n = p/2^{p+1}$.

Proof. See Section 7. $\square$

The following corollary is an easy consequence of Lemma 3.5.

Corollary 3.6. The random variable X which assigns to a type the number of its premises has the following distribution

$$X(p) = \begin{cases} 0 & \text{if } p = 0, \\ p/2^{p+1} & \text{if } p > 0, \end{cases}$$

expected value

$$E(X) = \sum_{p=1}^{\infty} p \frac{p}{2^{p+1}} = 3,$$

and variance

$$D^2(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2 = \sum_{p=1}^{\infty} p^2 \frac{p}{2^{p+1}} - 9 = 4,$$

so the standard deviation of X is 2.

The already presented properties concern the number of bracketings of a type. It is clear that the number of types of length n over a k-letter alphabet can be obtained by multiplying the number of bracketings by the number of possible letter patterns. In other words, the number of types of length n over an alphabet with k atomic types is

$$F_n^k = k^n C_n,$$

and the number of types with p premises of length $n \geq 2$ over an alphabet with k atomic types is

$$F_n^k(p) = k^n C_n(p).$$

It can be easily seen that $F_n^k = F_n^k(1) + \cdots + F_n^k(n-1)$, since there are no formulas of length n with more than $n-1$ premises. From Lemma 3.4 we can see that

$$F_n^k(1) = kF_{n-1}^k, \tag{12}$$

$$F_n^k(p) = \sum_{i=1}^{n-p} F_i^k F_{n-i}^k(p-1). \tag{13}$$

4. Types with one ground type

We start the discussion about the density of inhabited types from the types built from just one atomic type α. As noted by Statman in (Statman 1980) an implicational formula over one variable is an intuitionistic tautology iff it is a classical tautology. The simple proof of that fact which allows us to count inhabited types and then find their asymptotic density is given here in Lemma 4.3. In (Statman 1980, p. 529), Statman examines the probability of inhabitation, proving that it must lie between 0.625 and 0.86. We are going to prove that indeed the asymptotic probability exists and is equal to $1/2+\sqrt{5}/10$, which is almost exactly in the middle between the bounds of Statman.

Lemma 4.1. τ is a tautology iff $\tau \to \alpha$ is not a tautology, for τ over one variable α.

Proof. $(\to)$ Since τ is a tautology, τ is a classical tautology (see Lemma 2.2). So $\tau \to \alpha$ is not a classical tautology because of valuations which assign 0 to α. Again from Lemma 2.2, it is not a tautology.
$(\leftarrow)$ By induction on the length of τ. Suppose τ is not a tautology. Let $\tau = \tau[1], \ldots, \tau[n] \to \alpha$. First we show that all $\tau[i]$ for $1 \leq i \leq n$ are tautologies. If some $\tau[i]$ is not a tautology then by induction $\tau[i] \to \alpha$ is a tautology so $\tau = \tau[1], \ldots, \tau[n] \to \alpha$ is also a tautology; contradiction. So let $T_1 : \tau[1] \ldots T_n : \tau[n]$ be closed terms. Therefore $\lambda x.xT_1 \ldots T_n$ is a closed term of the type $(\tau[1], \ldots, \tau[n] \to \alpha) \to \alpha$ which ends the proof. □

Lemma 4.2. $\tau[1], \ldots, \tau[n] \to \alpha$ is not a tautology iff all $\tau[i]$ are tautologies, for $\tau[1], \ldots, \tau[n]$ over one variable α.

Proof. $(\to)$ By the same classical logic argument as in Lemma 4.1. $(\leftarrow)$ Assume some $\tau[i]$ is not a tautology. So $\tau[i] \to \alpha$ is a tautology and therefore $\tau[1], \ldots, \tau[n] \to \alpha$ must be also a tautology. □

Lemma 4.3. $\tau \to \mu$ is not a tautology iff τ is a tautology and μ is not a tautology, for τ and μ over one variable.

Proof. $(\to)$ Let $\mu = \mu[1] \ldots \mu[n] \to \alpha$. So the type $\tau \to \mu$ is of the form $\tau, \mu[1] \ldots \mu[n] \to \alpha$. By Theorem 4.2 all $\mu[i]$ and τ are tautologies. Therefore $\mu[1] \ldots \mu[n] \to \alpha$ is not a tautology.
$(\leftarrow)$ Follows from Theorem 4.2. □

Definition 4.4. By N^1_n and T^1_n we mean the number of noninhabited types (nontautologies) and inhabited types (tautologies), respectively, of length n built with one ground type.

Lemma 4.5. Numbers N^1_n of noninhabited types (nontautologies) of length n and T^1_n of inhabited types (tautologies) of length n are given by formulas:

$$N^1_1 = 1 \qquad\qquad N^1_n = \sum_{i=1}^{n-1} T^1_i N^1_{n-i},$$

$$T^1_1 = 0 \qquad\qquad T^1_n = F^1_n - \sum_{i=1}^{n-1} T^1_i N^1_{n-i}.$$

Proof. This is a simple consequence of Lemma 4.3. Every type τ of length n has the form $\tau_1 \to \tau_2$, where length of τ_1 is i and length of τ_2 is $n - i$. Since τ is a nontautology if and only if τ_1 is a tautology and τ_2 is a nontautology. Therefore the total number of nontautologies of length n is $\sum_{i=1}^{n-1} T_i^1 N_{n-i}^1$. □

The main theorem for this section is that the sequence T_n^1/F_n^1 is convergent. The meaning of this is the following: the asymptotic probability of the fact that a given type is inhabited (or that a given implicational formula is a tautology) exists and is equal to $1/2+\sqrt{5}/10$.

Theorem 4.6. The limit of the sequence T_n^1/F_n^1 exists and is equal to $1/2+\sqrt{5}/10$.

Proof. See Section 7. □

In the sequel we are going to consider the class of simple tautologies which are an important fragment of the set of tautologies.

Definition 4.7. A *simple tautology* (over an arbitrary alphabet) is a type $\tau = \tau_1, \ldots, \tau_n \to \alpha$ such that there is at least one component τ_i identical with α.

Evidently, a simple tautology is a tautology with a proof being the projection $\lambda x_1 \ldots x_n . x_i$. Let G_n^1 be the number of simple tautologies of length n built with one ground-type variable. In the next theorem we will find how big asymptotically is the fragment of simple one-letter tautologies within the set of all one-letter tautologies.

Theorem 4.8. Asymptotically $(25-5\sqrt{5})/18$ of all one letter tautologies are simple and $(5\sqrt{5}-7)/18$ are not simple.

Proof. The sequence G_n^1/T_n^1 is convergent and the limit is $(25-5\sqrt{5})/18$ because

$$\lim_{n\to\infty} G_n^1/T_n^1 = \frac{\lim_{n\to\infty} G_n^1/F_n^1}{\lim_{n\to\infty} T_n^1/F_n^1}.$$

See limits in Theorems 4.6 and 6.3 for $k=1$. □

5. Counting two-letter types

In this section we are going to find lower and upper bounds for asymptotic probabilities for the fraction of inhabited types of the types built with two letters. We will see that the upper bound for the fraction of tautologies (inhabited types), i.e. the fraction T_n^2/F_n^2, is smaller then the limit of T_n^1/F_n^1. This proves that asymptotically there are fewer tautologies (inhabited types) for two ground types than for one ground type.

By T_n^2 and N_n^2 we mean the numbers of tautologies and nontautologies, respectively, of length n over a two-letter alphabet. Of course $T_n^2 + N_n^2 = F_n^2$.

Definition 5.1. By the projection of a type τ we mean the one-letter type $\bar{\tau}$ obtained from τ by substituting all occurrences of ground-type letters in τ by one letter. We will employ the same notation for terms; so by $\bar{T}$ we mean the same term T with all typing information in T changed from type τ into $\bar{\tau}$.

Lemma 5.2. The projection of a tautology is a tautology.

Proof. If there is a closed typed lambda term T of type τ then the term $\bar{T}$ has a type $\bar{\tau}$. $\square$

Notice that Lemma 5.2 holds in general for tautologies built with any number of atomic types.

Definition 5.3. By ${}^{\otimes}N_n^2$ we mean the number of such two-letter nontautologies whose projections are one-letter tautologies, and ${}^{\downarrow}N_n^2$ means the number of such two-letters nontautologies whose projections are again one-letter nontautologies. Because classes of types described by numbers ${}^{\downarrow}N_n^2$ and ${}^{\otimes}N_n^2$ are disjoint, we have ${}^{\downarrow}N_n^2 + {}^{\otimes}N_n^2 = N_n^2$.

In the next theorem we will find the class of formulas which are nontautologies but their projections are tautologies. Moreover, the asymptotic density of the class is big enough to show that the total amount of nontautologies for two-letter types is bigger then for one-letter ones.

Lemma 5.4. For every two-letter tautology τ each formula in one of the forms

$$\begin{array}{ll} \tau, \alpha \to \beta & \qquad \tau, \beta \to \alpha \\ \alpha, \tau \to \beta & \qquad \beta, \tau \to \alpha \end{array}$$

is a two-letter nontautology while its projection is a tautology. Therefore ${}^{\otimes}N_n^2 \geq 4T_{n-2}^2$ for $n \geq 4$.

Proof. It is easy to notice that all those formulas are nontautologies since even classically they are not valid (see Lemma 2.2). Each projection must be a tautology since letters α and β are getting unified. The inequality holds since all four categories are disjoint for $n \geq 4$. $\square$

Lemma 5.5. For every $\varepsilon > 0$ there exists n_0 such that for every $n > n_0$ ${}^{\downarrow}N_n^2 \geq (1/2 - \sqrt{5}/10 - \varepsilon)F_n^2$.

Proof. It follows from Lemma 5.2 since every formula whose projection is a nontautology must be a nontautology. Therefore ${}^{\downarrow}N_n^2 = 2^n N_n^1$. Since N_n^1/F_n^1 is converging to the limit $1/2 - \sqrt{5}/10$ which is complementary to $1/2 + \sqrt{5}/10$ (see Theorem 4.6), then for any positive ε there is n_0 such that for every $n > n_0$ $N_n^1 \geq (1/2 - \sqrt{5}/10 - \varepsilon)F_n^1$. So ${}^{\downarrow}N_n^2 = 2^n N_n^1 \geq (1/2 - \sqrt{5}/10 - \varepsilon)2^n C_n = (1/2 - \sqrt{5}/10 - \varepsilon)F_n^2$. $\square$

Lemma 5.6. ${}^{\otimes}N_n^2 \geq \frac{1}{48}F_n^2$ for all sufficiently large n.

Proof. We use the definition of the class described by the number ${}^{\otimes}N_n^2$ and Lemma 5.4.

We have

$$\frac{{}^{\otimes}N_n^2}{F_n^2} \geq \frac{4T_{n-2}^2}{F_n^2} = \frac{F_{n-2}^2}{F_n^2}\,\frac{4T_{n-2}^2}{F_{n-2}^2} = \frac{2^{n-2}C_{n-2}}{2^n C_n}\,\frac{4T_{n-2}^2}{F_{n-2}^2} = \frac{1}{4}\,\frac{C_{n-2}}{C_n}\,\frac{4T_{n-2}^2}{F_{n-2}^2} \geq \frac{1}{48},$$

since $C_{n-2}/C_n > 1/16$ for all n (see (10)) and also $T_{n-2}^2/F_{n-2}^2 \geq G_{n-1}^2/F_{n-1}^2 \to 9/25 > 1/3$ (see Theorem 6.3 with $k = 2$). □

Theorem 5.7. There are fewer two-letters tautologies than single letter tautologies.

$$\limsup_{n\to\infty} \frac{T_n^2}{F_n^2} < \lim_{n\to\infty} \frac{T_n^1}{F_n^1}.$$

Proof. $N_n^2 = {}^{\downarrow}N_n^2 + {}^{\otimes}N_n^2$. By Lemma 5.5 (with $\varepsilon = 1/96$) there exists n_0 such that for $n > n_0$ ${}^{\downarrow}N_n^2 \geq (1/2-\sqrt{5}/10-1/96)F_n^2$. At the same time, by Lemma 5.6 , ${}^{\otimes}N_n^2 \geq F_n^2/48$ for $n \geq 4$. Therefore for $n > \max(n_0, 4)$ we get:

$$N_n^2 = {}^{\downarrow}N_n^2 + {}^{\otimes}N_n^2 \geq \left(\frac{1}{2} - \frac{\sqrt{5}}{10} - \frac{1}{96}\right) F_n^2 + \frac{1}{48}F_n^2 = \left(\frac{1}{2} - \frac{\sqrt{5}}{10} + \frac{1}{96}\right) F_n^2.$$

Consequently

$$\limsup_{n\to\infty} \frac{T_n^2}{F_n^2} \leq \frac{1}{2} + \frac{\sqrt{5}}{10} - \frac{1}{96} < \frac{1}{2} + \frac{\sqrt{5}}{10} = \lim_{n\to\infty} \frac{T_n^1}{F_n^1}$$

which completes the proof. □

Note that the general result proven below (Corollary 6.10) gives for $k = 2$ an upper bound $7/9 > 1/2 + \sqrt{5}/10$.

6. Types built with k atomic types

In this section we would like to investigate the asymptotic probability of finding inhabited types among types built with k variables. We will see that for any fixed k this probability is positive but the limit of those probabilities as $k \to \infty$ is zero.

By T_n^k and N_n^k we mean the numbers of intuitionistic tautologies and nontautologies, respectively, of length n over a k-letter alphabet. Of course $T_n^k + N_n^k = F_n^k$ and

$$F_n^k = \sum_{i=1}^{n-1} F_i^k\, F_{n-i}^k. \tag{14}$$

Definition 6.1. By a simple tautology we mean a type $\tau = \tau_1, \ldots, \tau_p \to \alpha$ for some ground type α, such that there is at least one component τ_i identical with α. Obviously, simple tautology is a tautology with a proof being a projection $\lambda x_1 \ldots x_p.x_i$. Let G_n^k be a number of simple tautologies of length n built with k ground type variables.

Lemma 6.2. The number G_n^k of simple tautologies is given the recursion

$$G_1^k = 0,$$
$$G_2^k = k,$$
$$G_n^k = \sum_{i=2}^{n-1} F_{n-i}^k G_i^k + (F_{n-1}^k - G_{n-1}^k).$$

Proof. The proof is based on two observations: First, $\tau_1 \to \tau_2$ is simple if τ_2 is simple. So for every type τ_1 of length $n-i$ and every simple tautology τ_2 of length i we have one simple tautology $\tau_1 \to \tau_2$ of length n. The sum starts from $i = 2$ because there are no simple tautologies of length 1. This part is responsible for the component $\sum_{i=2}^{n-1} F_{n-i}^k G_i^k$. The only other simple tautologies are those for which τ_1 is a ground type equal to the ground type τ_2 points to. Therefore for every formula μ of length $n-1$ which is not a simple tautology (there are exactly $F_{n-1}^k - G_{n-1}^k$ such formulas) we have exactly one simple tautology $a \to \mu$ where α is a ground type the type μ proves. □

Now we present a theorem about the asymptotic probability of the fact that a given type is a simple tautology, i.e. it has a proof which is a projection.

Theorem 6.3. The limit of the sequence G_n^k/F_n^k exists and is equal to $(4k+1)/(2k+1)^2$.

Proof. See Section 7 □

Corollary 6.4. For every $k \geq 1$ the asymptotic probability of finding an inhabited type in the set of all types built with k variables is positive, i.e. for any fixed $k \geq 1$ $\liminf_{n\to\infty} T_n^k/F_n^k > 0$.

Proof. For $k = 1$ the result was obtained in Theorem 4.6:

$$\liminf_{n\to\infty} \frac{T_n^1}{F_n^1} = \frac{1}{2} + \frac{\sqrt{5}}{10}.$$

For $k > 1$ the set of simple tautologies happens to be big enough to show the result. From Theorem 6.3 it follows that for all k:

$$\liminf_{n\to\infty} \frac{T_n^k}{F_n^k} \geq \lim_{n\to\infty} \frac{G_n^k}{F_n^k} = \frac{4k+1}{(2k+1)^2} > 0.$$

□

Definition 6.5. By "α-type" for some ground type α, we mean any type of the form $\tau_1, \ldots, \tau_n \to \alpha$. The type α is also an "α-type" with $n = 0$.

By a simple nontautology we mean an "α-type" τ such that all components of τ are not "α-types", for any ground type α.

It is quite obvious that a simple nontautology is neither an intuitionistic nor a classical tautology. Suppose the type $\tau_1, \ldots, \tau_n \to \alpha$ is a simple nontautology. Thus all τ_i are

"α_i-types" where $\alpha_i \neq \alpha$ for all $i \leq n$. Just put the 0-1 valuation to τ which evaluates all α_i to 1 and α to 0. Therefore τ is not classically valid and by Lemma 2.2 τ is not inhabited (τ is not an intuitionistic tautology).

Lemma 6.6. The number of simple nontautologies of length n built with k ground-type variables with p premises is

$$SN_n^k(p) = \left(\frac{k-1}{k}\right)^p F_n^k(p). \tag{15}$$

Proof. Let us consider types of length n and with p premises over k ground types, which have the form

$$(\tau_1 \to \alpha_1), (\tau_2 \to \alpha_2), \ldots, (\tau_p \to \alpha_p) \to \alpha_{p+1}, \tag{16}$$

for some fixed sequence of ground types $\alpha_1, \alpha_2, \ldots, \alpha_p, \alpha_{p+1}$. We permit any of the τ_i to be empty, and understand that in such cases $\tau_i \to \alpha_i$ is α_i. There are $F_n^k(p)/k^{p+1}$ such types because for each fixed sequence of ground types the number of types of the form (16) is clearly the same, every type of length n with p premises (there are $F_n^k(p)$ such types) is of the form (16) for some sequence $\alpha_1, \alpha_2, \ldots, \alpha_p, \alpha_{p+1}$, and there are k^{p+1} such sequences.
There are exactly $k(k-1)^p$ sequences $\alpha_1, \alpha_2, \ldots, \alpha_p, \alpha_{p+1}$, for which (16) is a simple nontautology. Therefore there are

$$\frac{F_n^k(p)}{k^{p+1}} k(k-1)^p = \left(\frac{k-1}{k}\right)^p F_n^k(p)$$

simple nontautologies with p premises. □

As in the case of tautologies, the number of simple nontautologies of the length n built with k ground type variables is

$$SN_n^k = \sum_{p=1}^{n-1} SN_n^k(p).$$

To calculate $\lim_{n\to\infty}(SN_n^k/F_n^k)$ let us first prove the auxiliary lemma:

Lemma 6.7.

$$\lim_{n\to\infty} \sum_{p=1}^{n-1} \left(\frac{k-1}{k}\right)^p \frac{C_n(p)}{C_n} = \sum_{p=1}^{\infty} \left(\frac{k-1}{k}\right)^p \frac{p}{2^{p+1}} = \frac{k(k-1)}{(k+1)^2}.$$

Proof. Consider the series

$$\sum_{p=1}^{\infty} \left(\frac{k-1}{k}\right)^p \frac{C_n(p)}{C_n}. \tag{17}$$

For all n the quotients $C_n(p)/C_n$ are bounded above by 1. Thus for any term of (17) we have the inequality

$$\left(\frac{k-1}{k}\right)^p \frac{C_n(p)}{C_n} \leq \left(\frac{k-1}{k}\right)^p.$$

The series $\sum_{p=1}^{\infty}((k-1)/k)^p$ converges, thus (17) converges uniformly in N. Since for a uniformly convergent series, the limit of its sum exists and is equal to the sum of the series of the limits of the elements, we can calculate the limit of (17) as follows:

$$\begin{aligned}
\lim_{n\to\infty} \sum_{p=1}^{n-1} \left(\frac{k-1}{k}\right)^p \frac{C_n(p)}{C_n} &= \sum_{p=1}^{\infty} \lim_{n\to\infty} \left(\frac{k-1}{k}\right)^p \frac{C_n(p)}{C_n} \quad \text{(because } C_n(p)=0 \text{ for } p \geq n\text{)} \\
&= \sum_{p=1}^{\infty} \left(\frac{k-1}{k}\right)^p \frac{p}{2^{p+1}} \qquad \text{(by Lemma 3.5)} \\
&= \frac{k(k-1)}{(k+1)^2}.
\end{aligned}$$

□

In Theorem 6.3 we have calculated the limit (as n tends to infinity) of the fraction of simple tautologies against all formulas of length n. Now we present the analogical theorem for simple nontautologies.

Theorem 6.8. The limit of the sequence SN_n^k/F_n^k exists and is equal to $k(k-1)/(k+1)^2$.

Proof. It is obvious that:

$$\begin{aligned}
SN_n^k(p) &= \left(\frac{k-1}{k}\right)^p F_n^k(p) \qquad \text{(by Lemma 6.6)} \\
&= \left(\frac{k-1}{k}\right)^p k^n C_n(p) \\
SN_n^k &= \sum_{p=1}^{n-1} \left(\frac{k-1}{k}\right)^p k^n C_n(p).
\end{aligned}$$

Thus

$$\begin{aligned}
\lim_{n\to\infty} \frac{SN_n^k}{F_n^k} &= \lim_{n\to\infty} \frac{\sum_{p=1}^{n-1} \left(\frac{k-1}{k}\right)^p k^n C_n(p)}{k^n C_n} \\
&= \lim_{n\to\infty} \sum_{p=1}^{n-1} \left(\frac{k-1}{k}\right)^p \frac{C_n(p)}{C_n}
\end{aligned}$$

$$= \sum_{p=1}^{\infty} \left(\frac{k-1}{k}\right)^p \frac{p}{2^{p+1}} \qquad \text{(by Lemma 6.7)}$$
$$= \frac{k(k-1)}{(k+1)^2}.$$

□

Corollary 6.9.

$$\lim_{k\to\infty} \lim_{n\to\infty} \frac{SN_n^k}{F_n^k} = 1.$$

It is quite interesting to note that the number of simple nontautologies is growing as k tends to infinity and asymptotically, all formulas are simple nontautologies.

Corollary 6.10. For all $k \geq 1$,

$$\limsup_{n\to\infty} \frac{T_n^k}{F_n^k} \leq \frac{3k+1}{(k+1)^2}.$$

Proof. Since $T_n^k = F_n^k - N_n^k$ and $N_n^k > SN_n^k$, thus we can calculate the limit of T_n^k/F_n^k as follows:

$$\begin{aligned}
\limsup_{n\to\infty} \frac{T_n^k}{F_n^k} &= \limsup_{n\to\infty} \frac{F_n^k - N_n^k}{F_n^k} \\
&\leq 1 - \lim_{n\to\infty} \frac{SN_n^k}{F_n^k} \\
&= 1 - \frac{k(k-1)}{(k+1)^2} \qquad \text{(by Theorem 6.8)} \\
&= \frac{3k+1}{(k+1)^2}.
\end{aligned}$$

□

7. Types and generating functions

In this section we have gathered three proofs making substantial use of generating functions.

7.1. *Calculating generating functions*

We start by calculating the generating functions we will need.

Lemma 7.1. The generating function f_F for the numbers F_n^k is

$$f_F(z) = \frac{1}{2} - \frac{1}{2}\sqrt{1 - 4\,k\,z},$$

and the generating function f_G for the numbers G_n^k is

$$f_G(z) = -\frac{-\frac{1}{2}z + \frac{1}{2}\,z\sqrt{1-4\,k\,z}}{\frac{1}{2} + \frac{1}{2}\sqrt{1-4\,k\,z} + z}.$$

Proof. Denote $f_F(z) = \sum_{n=1}^{\infty} F_n^k z^n$ and $f_G(z) = \sum_{n=1}^{\infty} G_n^k z^n$.
The recurrence (14)

$$F_n^k = \sum_{i=1}^{n-1} F_i^k F_{n-i}^k$$

becomes, after a closer examination, the equality

$$f_F \cdot f_F = f_F - k \cdot z \tag{18}$$

since the recurrence exactly corresponds to the multiplication of power series, with the exception that the linear term kz disappears.
Similarly, the recurrence

$$G_n^k = \sum_{i=2}^{n-1} F_{n-i}^k G_i^k + (F_{n-1}^k - G_{n-1}^k)$$

from Lemma 6.2 becomes

$$f_G = f_G \cdot f_F + z \cdot f_F - z \cdot f_G. \tag{19}$$

Solving (18) and (19) with the obvious boundary condition $f_F(0) = 0$ gives

$$f_F(z) = \frac{1}{2} - \frac{1}{2}\sqrt{1 - 4\,k\,z}$$

and

$$f_G(z) = -\frac{-\frac{1}{2}z + \frac{1}{2}\,z\sqrt{1-4\,k\,z}}{\frac{1}{2} + \frac{1}{2}\sqrt{1-4\,k\,z} + z}.$$

□

Corollary 7.2. The generating function for $C_n = F_n^1$ is

$$f_C(z) = \frac{1}{2} - \frac{1}{2}\sqrt{1-4z}.$$

□

Lemma 7.3. For fixed p the generating function for $C_n(p)$ is

$$f_p(z) = z(f_C(z))^p = z\left(\frac{1-\sqrt{1-4z}}{2}\right)^p.$$

Proof. Let $f_p(z)$ be a generating function for $C_n(p)$. Lemma 3.4 gives $C_n(p) = \sum_{i=1}^{n-p} C_i C_{n-i}(p-1)$ which becomes after a closer examination, the equality $f_{p-1}(z) \cdot f_C(z) = f_p(z)$. Since $C_n(1) = C_{n-1}$ we get $f_1(z) = z(f_C(z))$ and consequently $f_p(z) = z(f_C(z))^p$. □

Lemma 7.4. The generating function f_T for the numbers T_n^1 is

$$f_T(z) = \frac{3}{4} - \frac{1}{4}\sqrt{1-4\,z} - \frac{1}{4}\sqrt{2+2\sqrt{1-4\,z}+12\,z}.$$

Proof. The recurrence for N_n^1 and T_n^1 from Lemma 4.5 gives the following two identities for their generating functions $f_T(z) = \sum_{n=1}^{\infty} T_n^1 z^n$ and $f_N(z) = \sum_{n=1}^{\infty} N_n^1 z^n$.

$$\begin{aligned} f_N &= f_N(f_C - f_N) + z \\ f_T &= f_C - f_N. \end{aligned}$$

Solving with boundary condition $f_N(0) = 0$ we get

$$\begin{aligned} f_N(z) &= -\frac{1}{4} - \frac{1}{4}\sqrt{1-4\,z} + \frac{1}{4}\sqrt{2+2\sqrt{1-4\,z}+12\,z} \\ f_T(z) &= \frac{3}{4} - \frac{1}{4}\sqrt{1-4\,z} - \frac{1}{4}\sqrt{2+2\sqrt{1-4\,z}+12\,z}, \end{aligned}$$

as desired. □

7.2. *Calculation of the limits*

Here we give the three proofs of the results from the previous sections. They are ordered here not by their order of appearance in the paper, but rather by their difficulty.

We will frequently need the following small lemma.

Lemma 7.5.

$$\left|\binom{1/2}{n}\right| \sim c \cdot n^{-3/2}.$$

Proof. The generating function for C_n is $f_C(z) = \frac{1}{2} - \frac{1}{2}\sqrt{1-4z}$, which after substitution

$t = 4z$ becomes $\frac{1}{2} - \frac{1}{2}\sqrt{1-t}$. Directly from Corollary 2.4 we get

$$\begin{aligned} C_n &= \left(-\frac{1}{2}(-1)^n \binom{1/2}{n} + O(n^{-2})\right) 4^n \\ &= \left(-\frac{1}{2}(-1)^n \binom{1/2}{n} + O(n^{-2})\right) 4^n. \end{aligned} \tag{20}$$

If we compare that with (8) and use Stirling's Formula bellow

$$n! = (1 + o(1)) n^n e^{-n} \sqrt{2\pi n}.$$

the result is immediate. □

Lemma 3.5.

$$\lim_{n\to\infty} C_n(p)/C_n = p/2^{p+1}.$$

Proof. By Lemma 7.3, the generating function for $C_n(p)$ is

$$f_p(z) = z\left(\frac{1-\sqrt{1-4z}}{2}\right)^p.$$

Let $f = f_p(w/4) = (w/4)(1/2 - \sqrt{1-w}/2)^p$.
By the Newton Binomial Formula

$$4f/w = \sum_{j=0}^{p} \binom{p}{j} \frac{(-1)^j}{2^p} (1-w)^{p/2}.$$

The only singularity of the above function is at $w = 1$, hence by Corollary 2.4 we get

$$[w^n]\{4f/w\} = -\binom{p}{1} \frac{(-1)^n}{2^p} \binom{1/2}{n} + O(n^{-2}).$$

Therefore

$$\begin{aligned} C_n(p) &= [z^n]\{f_p\} \\ &= -\left(\frac{p(-1)^{n-1}}{2^p} \binom{1/2}{n-1} + O(n^{-2})\right) 4^{n-1} \\ &= -\left(\frac{p(-1)^{n-1}}{2^p} \binom{1/2}{n} \frac{n}{1/2-(n-1)} + O(n^{-2})\right) 4^{n-1} \\ &= -\left(\frac{p(-1)^{n-1}}{2^p} \binom{1/2}{n} (-1 + O(n^{-1})) + O(n^{-2})\right) 4^{n-1} \\ &= \left(\frac{p(-1)^{n-1}}{2^p} \binom{1/2}{n} + O(n^{-5/2}) + O(n^{-2})\right) 4^{n-1} \qquad \text{by Lemma 7.5} \\ &= \left(\frac{p(-1)^{n-1}}{2^p} \binom{1/2}{n} + O(n^{-2})\right) 4^{n-1}. \end{aligned}$$

We use formula (20) for calculating the quotient $C_n(p)/C_n$.

$$\begin{aligned}\frac{C_n(p)}{C_n} &= \frac{\left(\frac{p(-1)^{n-1}}{2^p}\binom{1/2}{n} + O(n^{-2})\right)4^{n-1}}{\left(\frac{(-1)^{n-1}}{2}\binom{1/2}{n} + O(n^{-2})\right)4^n} \\ &= \frac{\frac{p}{2^p}\binom{1/2}{n} + O(n^{-2})}{2\binom{1/2}{n} + O(n^{-2})} \\ &\sim \frac{p}{2^{p+1}}\end{aligned}$$

because $\left|\binom{1/2}{n}\right| \sim cn^{-3/2}$ by Lemma 7.5 □

Theorem 6.3.

$$\lim_{n\to\infty} G_n^k/F_n^k = \frac{4k+1}{(2k+1)^2}.$$

Proof. Recall that by Lemma 7.1

$$f_G(z) = -\frac{-\frac{1}{2}z + \frac{1}{2}z\sqrt{1-4kz}}{\frac{1}{2} + \frac{1}{2}\sqrt{1-4kz} + z}.$$

Let $t = 4kz$. After cancellation we get

$$f_G(t/4k) = -\frac{1}{2}\,\frac{t(\sqrt{1-t}-1)}{2k + 2k\sqrt{1-t} + t}.$$

Representing $f_G(t/4k)$ in terms of $\sqrt{1-t}$ gives

$$f_G(t/4k) = \frac{1}{2}\cdot\frac{(\sqrt{1-t})^2 - 2\sqrt{1-t} + 1}{-\sqrt{1-t} + 2k + 1} = -k + \frac{1}{2} - \frac{1}{2}\sqrt{1-t} + 2\frac{k^2}{-\sqrt{1-t} + 2k + 1}.$$

This representation reveals that the only singularity of $f_G(t/4k)$ located in $|t| \le 1$ is indeed $t = 1$. The other singularity, located where the denominator in the last fraction becomes 0, is achieved when $\sqrt{1-t} = 2k+1$, i.e., for t of a modulus substantially greater than 1. So we can use Corollary 2.4.
The expansion of $f_G(t/4k)$ near $t = 1$ is

$$f_G(t) = \left(-k + \frac{1}{2} + \frac{2k^2}{2k+1}\right) + \left(-\frac{1}{2} + \frac{2k^2}{(2k+1)^2}\right)\sqrt{1-t} + \ldots$$

Therefore by Corollary 2.4 we get

$$[t^n]\{f_G(t/4k)\} = \left(-\frac{1}{2} + \frac{2k^2}{(2k+1)^2}\right)(-1)^n\binom{1/2}{n} + O(n^{-2}),$$

hence

$$G_n^k = [z^n]\{f_G(z)\} = \left(\left(-\frac{1}{2} + \frac{2k^2}{(2k+1)^2}\right)(-1)^n \binom{1/2}{n} + O(n^{-2})\right)(4k)^n.$$

Calculating the quotient G_n^k / F_n^k and taking into account that $F_n^k = \left(2(-1)^{n-1}\binom{1/2}{n} + O(n^{-2})\right) 4^{n-1} k^n$ by (20) we get

$$\begin{aligned} \frac{G_n^k}{F_n^k} &= \frac{((-1/2 + 2k^2/(2k+1)^2)(-1)^n \binom{1/2}{n} + O(n^{-2}))(4k)^n}{(2(-1)^{n-1}\binom{1/2}{n} + O(n^{-2}))4^{n-1}k^n} \\ &= \frac{1/2 - 2k^2/(2k+1)^2}{1/2}(1 + o(1)) \\ &= \left(1 - \frac{4k^2}{(2k+1)^2}\right)(1 + o(1)) \\ &\sim \frac{4k+1}{(2k+1)^2}, \end{aligned}$$

as desired. $\square$

Theorem 4.6.

$$\lim_{n\to\infty} T_n^1 / C_n = \frac{1}{2} + \frac{\sqrt{5}}{10}.$$

Proof. By Lemma 7.4,

$$f_T(z) = \frac{3}{4} - \frac{1}{4}\sqrt{1-4z} - \frac{1}{4}\sqrt{2 + 2\sqrt{1-4z} + 12z}.$$

With substitution $t = 4z$ we get

$$f_T(t/4) = \frac{3}{4} - \frac{1}{4}\sqrt{1-t} - \frac{1}{4}\sqrt{2 + 2\sqrt{1-t} + 3t}.$$

In order to use Szegö Lemma we must know the singularities of $f_T(t/4)$. Apart from $t = 1$ there can be only a singularity where the expression under the outer root sign becomes 0. This happens at $t = -16/9$, so indeed $t = 1$ is the only singularity of $f_T(t/4)$ in $|t| \le 1$.
Expressing f_T as a function of $\sqrt{1-t}$ we get

$$f_T(t/4) = \frac{3}{4} - \frac{1}{4}\sqrt{1-t} - \frac{1}{4}\sqrt{5 + 2\sqrt{1-t} - 3(\sqrt{1-t})^2},$$

whose expansion near $t = 1$ is

$$f_T(t/4) = \left(\frac{3}{4} - \frac{1}{4}\sqrt{5}\right) + \left(-\frac{1}{4} - \frac{1}{20}\sqrt{5}\right)\sqrt{1-t} + \ldots$$

By Corollary 2.4 we get

$$T_n^1 = \left(\left(-\frac{1}{4} - \frac{\sqrt{5}}{20}\right)(-1)^n \binom{1/2}{n} + O(n^{-2})\right) 4^n.$$

Again as in the previous proof we get

$$\begin{aligned}
\frac{T_n^1}{C_n} &= \frac{((-\frac{1}{4} - \frac{\sqrt{5}}{20})(-1)^n \binom{1/2}{n} + O(n^{-2})\,)4^n}{(\,2(-1)^{n-1}\binom{1/2}{n} + O(n^{-2})\,)4^{n-1}} \\
&= \frac{1 + 4\sqrt{5}/20}{2}(1 + o(1)) \\
&\sim \frac{1}{2} + \frac{\sqrt{5}}{10} \\
&= 0.7236067978\ldots
\end{aligned}$$

This concludes the proof. □

8. Conclusions and open problems

Let us sum up the presented results:

1 There exists an asymptotic probability for one-letter tautologies (one-letter inhabited types) $\lim_{n\to\infty} T_n^1/F_n^1 = 1/2 + \sqrt{5}/10$ (Theorem 4.6)

2 There are fewer two-letters tautologies (two-letters inhabited types) than one-letter tautologies $\limsup_{n\to\infty} T_n^2/F_n^2 < \lim_{n\to\infty} T_n^1/F_n^1$ (Theorem 5.7)

3 There exists an asymptotic probability for simple tautologies for any number of ground types $\lim_{n\to\infty} G_n^k/F_n^k = (4k+1)/(2k+1)^2$ (Theorem 6.3)

4 There exists an asymptotic probability for simple nontautologies for any number of ground types $\lim_{n\to\infty} SN_n^k/F_n^k = k(k-1)/(k+1)^2$ (Theorem 6.8)

5 Asymptotically $(4k+1)/(2k+1)^2 < T^k/F^k < (3k+1)/(k+1)^2$ (Corollaries 6.4 and 6.10).

As we can see, the gap between asymptotic lower and upper bounds for T_n^k/F_n^k is relatively small and asymptotically the upper bound is 3 times bigger then the lower bound. The main problem is to close this gap and prove the convergence of T_n^k/F_n^k for all $k \geq 2$. However the problem seems difficult since it requires a syntactic description of the notion of a tautology which is inherently semantical.

Notice that all results are valid for both classical and intuitionistic logic, i.e.,

$$\frac{4k+1}{(2k+1)^2} \leq \liminf_{n\to\infty} \frac{CL_n^k}{F_n^k} \leq \limsup_{n\to\infty} \frac{CL_n^k}{F_n^k} \leq \frac{3k+1}{(k+1)^2},$$

where CL_n^k is the number of classical tautologies of length n with k letters. This is because simple tautologies are classical tautologies, and simple nontautologies are neither intuitionistic nor classical tautologies. Thus we can calculate the fraction of intuitionistic tautologies among classical tautologies.

Theorem 8.1.

$$\lim_{k\to\infty} \liminf_{n\to\infty} \frac{T_n^k}{CL_n^k} \geq \frac{1}{3}.$$

Proof.

$$\begin{aligned}
\limsup_{n\to\infty} \frac{CL_n^k}{T_n^k} &\leq \limsup_{n\to\infty} \frac{CL_n^k}{G_n^k} \\
&= \limsup_{n\to\infty} \frac{CL_n^k}{F_n^k} \Big/ \lim_{n\to\infty} \frac{G_n^k}{F_n^k} \\
&\leq \frac{3k+1}{(k+1)^2} \Big/ \frac{4k+1}{4k^2+4k+1} \\
&\stackrel{k\to\infty}{\longrightarrow} 3.
\end{aligned}$$

Thus

$$\lim_{k\to\infty} \liminf_{n\to\infty} \frac{T_n^k}{CL_n^k} \geq \frac{1}{3}.$$

$\square$

The result $1/3$ has been obtained using very rough estimations. However, basing on our experiences of counting the intuitionistic tautologies, we strongly believe that the limit of T_n^k/CL_n^k exists and is equal to one. Although intuitionistic logic attracts the attention of logicians and computer scientists, it is often underestimated by classical mathematicians. Such a result would prove that a random classical tautology has an intuitionistic proof with the probability 1. Moreover, we believe that the proportion of simple tautologies to intuitionistic tautologies tends to one as the length of formulas tends to infinity. It would be an even more unexpected result, because it would show that, in fact, a randomly chosen proof is a projection and statistically all true statements are the trivial ones. At the moment, these are just expectations but certainly it is worth to look in this direction.

Conjecture 1. For every $k \geq 1$ the limit $\lim_{n\to\infty} T_n^k/CL_n^k$ exists and is equal to 1.

Conjecture 2. For every $k \geq 1$ the limit $\lim_{n\to\infty} G_n^k/T_n^k$ exists is equal to 1.

Acknowledgement. We are very grateful to the anonymous referee who suggested many improvements to the presentation of our paper. We are also grateful to Pawel Idziak for the valuable discussion on the early version of this paper.

References

Barendregt, H.P. (1984) The lambda calculus. Its syntax and semantics. Revised edition. *Studies in Logic and the Foundations of Mathematics*, **103**, North-Holland, Amsterdam-New York.

Comtet, L. (1974) Advanced combinatorics. The art of finite and infinite expansions. Revised and enlarged edition. Reidel, Dordrecht.

Hindley, J.R. (1997) Basic Simple Type Theory. *Cambridge Tracts in Theoretical Computer Science*, **42** Cambridge University Press.

Statman, R. (1980) On the existence of closed terms in the typed λ-calculus. In: Hindley, J.R., and Seldin, J., (eds.) *Combinatory Logic, Lambda Calculus and Formalism, Academic Press*, New York.

Szegö, G. (1975) Orthogonal polynomials. Fourth edition. *AMS, Colloquium Publications*, **23**, Providence.

Wolfram, D.A. (1993) The Clausual Theory of Types. *Cambridge Tracts in Theoretical Computer Science*, **21** Cambridge University Press.

Wilf, H.S. (1994) Generatingfunctionology. Second edition. *Academic Press*, Boston.