

How big is BCI fragment of BCK logic*

Katarzyna Grygiel, Paweł M. Idziak and Marek Zaionc†

September 24, 2018

Abstract

We investigate quantitative properties of BCI and BCK logics. The first part of the paper compares the number of formulas provable in BCI versus BCK logics. We consider formulas built on implication and a fixed set of k variables. We investigate the proportion between the number of such formulas of a given length n provable in BCI logic against the number of formulas of length n provable in richer BCK logic. We examine an asymptotic behavior of this fraction when length n of formulas tends to infinity. This limit gives a probability measure that randomly chosen BCK formula is also provable in BCI. We prove that this probability tends to zero as the number of variables tends to infinity. The second part of the paper is devoted to the number of lambda terms representing proofs of BCI and BCK logics. We build a proportion between number of such proofs of the same length n and we investigate asymptotic behavior of this proportion when length of proofs tends to infinity. We demonstrate that with probability 0 a randomly chosen BCK proof is also a proof of a BCI formula.

Keywords: BCK and BCI logics, asymptotic probability in logic, analytic combinatorics.

1 Introduction

The results presented in this paper are a part of research in which the likelihood of truth is estimated for various propositional logics with a limited number of variables. Probabilistic methods appear to be very powerful in combinatorics and computer science. From a point of view of these methods we investigate a typical object chosen from some set. For formulas in the fixed propositional language, we investigate the proportion between the number of valid formulas of a given length n against the number of all formulas of length n . Our interest lies in finding the limit of that fraction when n tends to infinity. If the limit exists, then it is represented by a real number which we may call *the density* of the investigated logic. In general, we are also interested in finding the ‘*density*’ of some other classes of formulas. Good presentation and overview of asymptotic methods for random boolean expressions can be found in the paper [9] of Gardy. For the purely implicational logic of one variable (and at the same time simple type systems), the exact value of the density of true formulas was computed by Moczurad, Tyszkiewicz and Zaionc in [21]. The classical logic of one variable and the two connectives of implication and negation was studied in Zaionc [30]; over the same language, the exact proportion between

*This work was partially supported by grant number N206 3761 37, Polish Ministry of Science and Higher Education.

†Faculty of Mathematics and Computer Science, Theoretical Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. Email: {grygiel,idziak,zaionc}@tcs.uj.edu.pl

intuitionistic and classical logics was determined by Kostrzycka and Zaionc in [15]. Asymptotic identity between classical and intuitionistic logic of implication has been proved in Fournier, Gardy, Genitrini and Zaionc in [7]. Some variants involving expressions with other logical connectives have also been considered. Genitrini and Kozik have studied the influence of adding the connectors \vee and \wedge to implication in [11], while Matecki in [20] considered the case of the single equivalence connector. For two connectives again, the *and/or* case has already received much attention – see Lefmann and Savický [19], Chauvin, Flajolet, Gardy and Gittenberger [3], Gardy and Woods [10], Woods [27] and Kozik [17]. Let us also mention the survey [9] of Gardy on the probability distributions on Boolean functions induced by random Boolean expressions; this survey deals with the whole set of Boolean functions on some finite number of variables.

2 BCK and BCI logics

The logics BCK and BCI are ones of several pure implication calculi. Its name comes from the connection with the combinators B, C, K and I (see [2]). From the perspective of type theory, BCK and BCI can be viewed as the set of types of a certain restricted family of lambda terms, via the Curry-Howard isomorphism. Formally, logics BCK and BCI can be defined, each one separately, as Hilbert systems by three axiom schemes and detachment rule. Namely BCK is based on B, C and K while BCI is based on B, C and I where:

- (B) $(\varphi \Rightarrow \psi) \Rightarrow ((\chi \Rightarrow \varphi) \Rightarrow (\chi \Rightarrow \psi))$ (prefixing)
- (C) $(\varphi \Rightarrow (\psi \Rightarrow \chi)) \Rightarrow (\psi \Rightarrow (\varphi \Rightarrow \chi))$ (commutation)
- (K) $\varphi \Rightarrow (\psi \Rightarrow \varphi)$
- (I) $\varphi \Rightarrow \varphi$ (identity)

In BCK we are able to prove I therefore the logic BCI is a subset of the BCK. Let us observe that implicational formulas may be seen as rooted binary trees.

Definition 1. *By a formula tree we mean the rooted binary tree in which nodes are labeled by \Rightarrow and have two successors left and right while leaves of the tree are labeled by variables.*

Definition 2. *With every implicational formula φ we associate the formula tree $G(\varphi)$ in the following way:*

- If x is a variable, then $G(x)$ is a single node labeled with x .
- Tree $G(\varphi \Rightarrow \psi)$ is the tree with the new root labeled with \Rightarrow and two subtrees: left $G(\varphi)$ and right $G(\psi)$.

3 λ -calculus as a proof system

Lambda calculus is a standard mechanism for proof system representation for various propositional calculi. By the Curry-Howard isomorphism there is a one-to-one correspondence between provable formulas in intuitionistic implicational logic and types of closed lambda calculus terms. Moreover, proofs of formulas correspond to typable terms. We start with presenting some fundamental concepts of the λ -calculus, as well as with some new definitions used in this paper.

Definition 3. *Let V be a countable set of variables. The set Λ of λ -terms is defined by the following grammar:*

1. every variable is a lambda term,
2. if t and s are lambda terms then ts is a lambda term,
3. if t is a lambda terms and x is a variable then $\lambda x.t$ is a lambda term.

As usual, λ -terms are considered modulo the α -equivalence, i.e. two terms which differ only by the names of bounded variables are considered equal. Observe that λ -terms can be seen as rooted unary-binary trees.

Definition 4. By a lambda tree we mean the following rooted graph with two kinds of edges: undirected and directed. One distinguished node is called the root of the graph. The graph induced by undirected edges is a rooted tree with the distinguished node being the root of it. There are two kinds of internal nodes labeled by $@$ and by λ . Nodes labeled by $@$ have two successors left and right. Nodes labeled with λ have only one successor. Leaves of the tree are either labeled by variables or are connected by directed edge with the one of λ nodes placed on the path from it to the root.

Definition 5. With every lambda term t we associate the lambda tree $G(t)$ in the following way:

- If x is a variable then $G(x)$ is a single node labeled with x .
- Lambda tree $G(PQ)$ is a lambda tree with the new root labeled with $@$ and connected by two new undirected edges with roots of two lambda subtrees left $G(P)$ and right $G(Q)$.
- Tree $G(\lambda x.P)$ is obtained from $G(P)$ in four steps:
 - Add new root node labeled with λ .
 - Connect new root by undirected edge with the root of $G(P)$.
 - Connect all leaves of $G(P)$ labeled with x by directed edges with the new root.
 - Remove all labels x from $G(P)$.

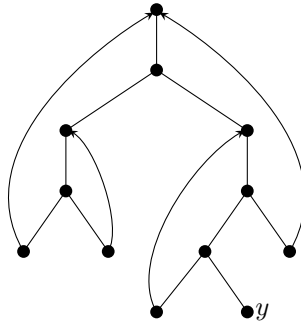


Figure 1: The lambda tree representing the term $\lambda z.(\lambda u.zu)((\lambda u.uy)z)$

Observation 6. If T is a lambda tree, then $T = G(M)$ for some lambda term M . Terms M and N are α -equivalent iff $G(M) = G(N)$. Free variables of term M are the same as variables labeling leafs of $G(M)$.

We often use (without giving the precise definition) the classical terminology about trees (e.g. path, root, leaf, etc.). A path from the root to a leaf is called a branch.

3.1 BCI and BCK classes of lambda terms

The Curry-Howard isomorphism for proof representation in entire intuitionistic implicational logic can be restricted to weaker logics. Therefore, we can look at the axioms (B), (C), (I) and (K) of BCI and BCK logics as at types à la Curry in the typed lambda calculus. Lambda terms for which these types are principal (the most general ones) are respectively ([13]):

$$\begin{aligned}\mathbf{B} &\equiv \lambda xyz.x(yz), \\ \mathbf{C} &\equiv \lambda xyz.xzy, \\ \mathbf{I} &\equiv \lambda x.x, \\ \mathbf{K} &\equiv \lambda xy.x.\end{aligned}$$

Theorem 7. *The combinator I is provable in BCK.*

Proof. For example the lambda term $(\mathbf{CK})\mathbf{K} = \lambda z.z$ forms proof for the combinator I in logic BCK. \square

Theorem 8. *BCI is a proper subset of BCK*

Proof. Inclusion follows from Theorem 7. Combinator K is not provable in logic BCI (see [1], for example). \square

We are going to isolate the special set of BCK provable formulas called simple tautologies which forms a simple and large fragment of the set of all BCK provable formulas. As we will see afterwards the class of simple tautologies is so big that it can play a role of good approximation of the whole set of BCK tautologies. Therefore, quantitative investigations about behavior of the whole set can be nicely approximated by this fragment. See [31] for discussion about quantitative aspects of simple tautologies.

Definition 9. *A simple tautology is an implicational formula of the form $\tau_1 \Rightarrow (\dots \Rightarrow (\tau_p \Rightarrow \alpha) \dots)$ such that $p > 0$, α is a variable and there is at least one component τ_i identical to α .*

Theorem 10. *Every simple tautology is BCK provable.*

Proof. By $(\mathbf{CK})\mathbf{K}$ we can prove $\alpha \Rightarrow \alpha$. Using several times axiom (K) we can add any number of premisses and prove $\tau_1 \Rightarrow (\dots \Rightarrow (\tau_p \Rightarrow (\alpha \Rightarrow \alpha)) \dots)$. Using axiom (C) we are able to permute premisses to get $\tau_1 \Rightarrow (\dots \Rightarrow (\tau_p \Rightarrow \alpha) \dots)$. \square

Definition 11. *The smallest class of lambda calculus terms containing \mathbf{B} , \mathbf{C} and \mathbf{I} (resp. \mathbf{B} , \mathbf{C} and \mathbf{K}) and closed under application and β -reduction is called the class of BCI (resp. BCK) lambda terms.*

Theorem 12. (1) *A BCI lambda term is a closed lambda term P such that*

- (i) *for each subterm $\lambda x.M$ of P , x occurs free in M exactly once,*
- (ii) *each free variable of P has just one free occurrence in P .*

(2) *A BCK lambda term is a closed lambda term P such that*

- (i) *for each subterm $\lambda x.M$ of P , x occurs free in M at most once,*
- (ii) *each free variable of P has just one free occurrence in P .*

Proof. Proof can be found in Roger Hindley's book [13]. \square

By Theorem 12 we immediately get that BCI is a proper subclass of BCK.

4 Classes of formulas

Definition 13. The language F^k over k propositional variables $\{a_1, \dots, a_k\}$ is defined inductively as:

$$\begin{aligned} a_i &\in F^k \quad \text{for all } i \leq k, \\ \phi \Rightarrow \psi &\in F^k \quad \text{if } \phi \in F^k \text{ and } \psi \in F^k. \end{aligned}$$

We can now define the usual notation for a formula. Let $T \in F^k$ be a formula. Hence it is of the form $A_1 \Rightarrow (A_2 \Rightarrow (\dots \Rightarrow (A_p \Rightarrow r(T)))) \dots$; we shall write it

$$T = A_1, \dots, A_p \Rightarrow r(T).$$

The formulas A_i are called the *premisses* of T and the rightmost propositional variable $r(T)$ of the formula is called the *goal* of T . For formula T which is itself a propositional variable obviously $p = 0$ and $r(T) = T$. To prove quantitative results about BCK and BCI logics we need to define several other classes of formulas, all of them being special kinds of either tautologies or non-tautologies.

Definition 14. We define the following subsets of F^k :

- The set of all classical tautologies, CL^k is the set of formulas which are true under any $\{0, 1\}$ valuation.
- The set of all intuitionistic tautologies, INT^k is the set of formulas for which there are closed lambda terms (constructive proofs) of type identical with the formula.
- The set of all Peirce formulas, $PEIRCE^k$ is the set of classical tautologies which are not intuitionistic ones.
- The set BCK^k is the set of formulas for which there are closed lambda BCK terms of type identical with the formula.
- The set BCI^k is the set of formulas for which there are closed lambda BCI terms of type identical with the formula.
- The set of simple tautologies, G^k is the set of expressions that can be written as

$$T = A_1, \dots, A_p \Rightarrow r(T),$$

where at least one of A_i 's is the variable $r(T)$.

- The set of even formulas $EVEN^k$, is the set of formulas in which each variable occurs even number of times.
- The set of simple non-tautologies SN^k , is the set of formulas of the form

$$T = A_1, \dots, A_p \Rightarrow r(T),$$

where $r(A_i) \neq r(T)$ for all i .

- The set LN^k is the set of less simple non-tautologies, defined as the set of formulas of the form

$$T = B_1, \dots, B_{i-1}, C, B_i, \dots, B_p \Rightarrow r(T),$$

such that

$$C = C_1, C_2, \dots, C_q \Rightarrow r(C),$$

where $r(C) = r(T)$, $q \geq 1$, and

$$C_1 = D_1, D_2, \dots, D_r \Rightarrow r(D),$$

where $r(D) \neq r(T)$, $r \geq 0$, and the following holds: for all j , $r(B_j) \notin \{r(T), r(D)\}$ and $r(D_j) \notin \{r(T), r(D)\}$.

The obvious relations between classes above are the following.

Lemma 15. $SN^k \cup LN^k \subseteq F^k \setminus CL^k$

Proof. Suppose $T = A_1, \dots, A_p \Rightarrow r(T)$ is in SN^k . Then evaluate, all of assumptions A_i by 1 and goal $r(T)$ by 0 we get that $T \notin CL$. Now let $T \in LN^k$ and T is in the form described by the definition 14. The shape of T allows us to evaluate $r(T)$ and $r(D)$ by 0 and all the $r(B_j)$ and $r(D_j)$ by 1 to see that $T \notin CL$. \square

Lemma 16. $SN^k \cap LN^k = \emptyset$

Proof. Simply by observing the syntactic structure of both sets. \square

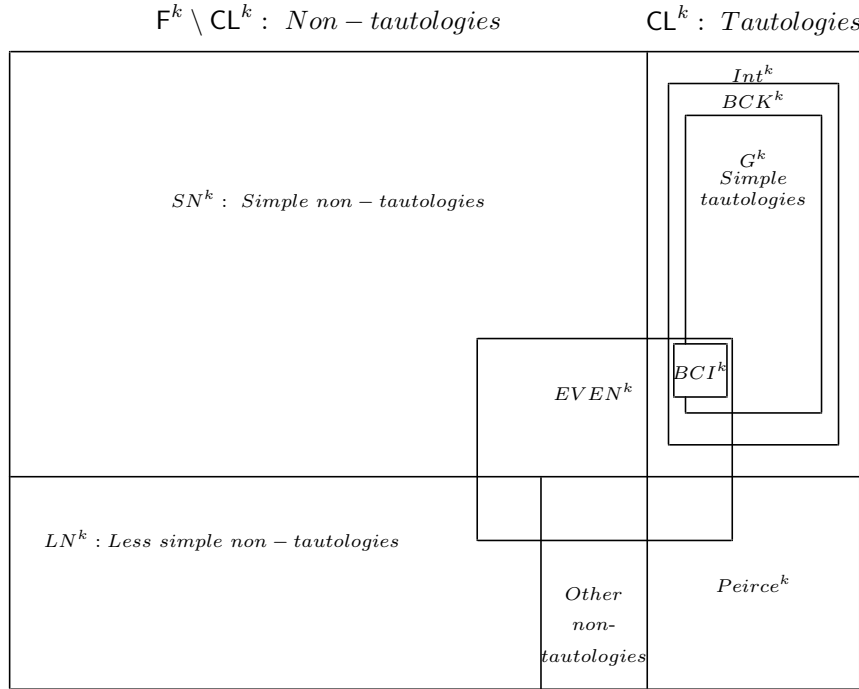


Figure 2: Inclusions summarized in Lemmas 15 till 18

Lemma 17. $G^k \subseteq BCK^k \subseteq INT^k \subsetneq CL^k \subsetneq F^k \setminus (SN^k \cup LN^k)$

Proof. The first inclusion can be found in Theorem 10. The rest is trivial. \square

Lemma 18. $BCI^k \subseteq EVEN^k \cap BCK^k$

Proof. $BCI^k \subsetneq EVEN^k$ has been proved recently by Tomasz Kowalski in the paper [16] Theorem 6.1. $BCI^k \subsetneq BCK^k$ is a classical fact which can be found in [1]. See also Theorem 8. \square

5 Densities of sets of formulas

First we establish the way in which the size of formula trees are measured.

Definition 19. By $\|\phi\|$ we mean the size of formula ϕ which we define as total number of leaves in the formula tree $G(\phi)$. This is in fact the total number of occurrences of propositional variables in the formula. Formally,

$$\|a_i\| = 1 \text{ and } \|\phi \Rightarrow \psi\| = \|\phi\| + \|\psi\|.$$

Definition 20. We associate the density $\mu(\mathbf{X})$ with a subset $\mathbf{X} \subseteq F^k$ of formulas as:

$$\mu(\mathbf{X}) = \lim_{n \rightarrow \infty} \frac{\#\{t \in \mathbf{X} : \|t\| = n\}}{\#\{t \in F^k : \|t\| = n\}} \quad (1)$$

if the limit exists.

The number $\mu(\mathbf{X})$ if it exists is an asymptotic probability of finding a formula from the class \mathbf{X} among all formulas from F^k or it can be interpreted as the asymptotic density of the set \mathbf{X} in the set F^k . It can be immediately seen that the density μ is finitely additive so if \mathbf{X} and \mathbf{Y} are disjoint classes of formulas such that $\mu(\mathbf{X})$ and $\mu(\mathbf{Y})$ exist then $\mu(\mathbf{X} \cup \mathbf{Y})$ also exists and $\mu(\mathbf{X} \cup \mathbf{Y}) = \mu(\mathbf{X}) + \mu(\mathbf{Y})$. It is straightforward to observe that for any finite set \mathbf{X} the density $\mu(\mathbf{X})$ exists and is 0. Dually for co-finite sets \mathbf{X} the density $\mu(\mathbf{X}) = 1$. The density μ is not countably additive so in general the formula

$$\mu\left(\bigcup_{i=0}^{\infty} \mathbf{X}_i\right) = \sum_{i=0}^{\infty} \mu(\mathbf{X}_i) \quad (2)$$

does not hold for all pairwise disjoint classes of sets $\{\mathbf{X}_i\}_{i \in \mathbb{N}}$. A good counterexample for the equation (2) is to take as \mathbf{X}_i the singleton of i -th formula from our language under any natural order of formulas. On the left hand side of equation 2 we get $\mu(F^k)$ which is 1 but on right hand side $\mu(\mathbf{X}_i) = 0$ for all $i \in \mathbb{N}$ and so the sum is 0. Finally, we define:

$$\begin{aligned} \mu^-(\mathbf{X}) &= \liminf_{n \rightarrow \infty} \frac{\#\{t \in \mathbf{X} : \|t\| = n\}}{\#\{t \in F^k : \|t\| = n\}} \\ \mu^+(\mathbf{X}) &= \limsup_{n \rightarrow \infty} \frac{\#\{t \in \mathbf{X} : \|t\| = n\}}{\#\{t \in F^k : \|t\| = n\}} \end{aligned}$$

These two numbers are well defined for any set of formulas \mathbf{X} , even when the limiting ratio $\mu(\mathbf{X})$ is not known to exist.

5.1 Enumerating formulas

In this section we present some properties of numbers characterizing the amount of formulas in different classes defined in our language. Many results and methods could be rephrased purely in terms of binary trees with given properties. Obviously an implicational formula from F^k of size n can be seen as a binary tree with n leaves and k labels per leaf (see definitions 1 and 2). We will analyze several classes of formulas (trees).

Definition 21. By F_n^k we mean the total number of formulas from F^k of size n so:

$$F_n^k = \#\{\phi \in F^k : \|\phi\| = n\}. \quad (3)$$

Lemma 22. *The number $F_n^k = k^n C_n$ where C_n is $(n-1)$ th Catalan number.*

Proof. We may use combinatorial observation. A formula from F^k of size n can be interpreted as full binary tree of n leaves with k label per leaf. Therefore for $n = 0$ and $n = 1$ it is obvious. Any formula of size $n > 1$ is the implication (tree) between some pair of formulas (trees) of sizes i and $n-i$, respectively. Therefore the total number of such pairs is $\sum_{i=1}^{n-1} F_i^k F_{n-i}^k$. Therefore by simple induction we can immediately see that $F_n^k = k^n C_n$. For more elaborate treatment of Catalan numbers see Wilf [26, pp. 43–44]. We mention only the following well-known nonrecursive formula for $C_n = \frac{1}{n} \binom{2n-2}{n-1}$. \square

Lemma 23. *The number G_n^k of simple tautologies is given by the recursion*

$$G_1^k = 0, \quad G_2^k = k, \quad (4)$$

$$G_n^k = F_{n-1}^k - G_{n-1}^k + \sum_{i=2}^{n-1} F_{n-i}^k G_i^k. \quad (5)$$

Proof. For the whole discussion about simple tautologies see [31]. In particular for a proof look at Lemma 15 of [31]. \square

Lemma 24. *The number $EVEN_n^k$ is given by:*

$$\frac{C_n}{2^k} \sum_{j=0}^k \binom{k}{j} (k-2j)^n.$$

Proof. The proof is based on the observation obtained in the paper of Franssens ([8] page 30, formula 7.18). In this paper it is obtained an explicit formula $e_n^k = \frac{1}{2^k} \sum_{j=0}^k \binom{k}{j} (k-2j)^n$ for the number of closed walks, based at a vertex, of length n along the edges of k -dimensional cube (see also [23]). This number in fact appears at Maclaurin series of $\cosh^k(t)$, for all k . Note that for odd n this gives 0. As we can see the number e_n^k obviously enumerates the set of all sequences of length n of variables $\{a_1, \dots, a_k\}$ in which every variable a_i occurs even number of times. Multiplying this by the number C_n of all binary trees with n leaves we obtain the explicit formula for $EVEN_n^k$. For some special k , sequences e_n^k are mentioned in Sloane's catalogue [22]. For instance, $e_{2n}^2 = 2 * 4^{n-2}$ is described in Sloane as A009117. The sequence $e_{2n}^3 = (3^n + 3)/4$ is present in Sloane's catalogue as A054879. Finally e_{2n}^4 is Sloane's A092812. \square

5.2 Generating functions

In this paper we investigate the proportion between the number of formulas of the size n that are tautologies in various logics against the number of all formulas of size n for propositional formulas of the language F^k . Our interest lies in finding limit of that fraction when n tends to infinity. For this purpose combinatorics has developed an extremely powerful tool, in the form of generating series and generating functions. A nice exposition of the method can be found in Wilf [26], as well as in Flajolet, Sedgewick [6]. As the reader may now expect, while working with formulas we will be often concerned with complex analysis, analytic functions and their singularities.

Let $A = (A_0, A_1, A_2, \dots)$ be a sequence of real numbers. 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 $\{z \in \mathbb{C} : |z| < R\}$ 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 neighborhood 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 neighborhood of z_0 , into a power series $\sum_{n=0}^{\infty} A_n(z - z_0)^n$ is unique.

Many questions concerning the asymptotic behavior of A can be efficiently resolved by analyzing the behavior of its generating function f_A at the complex circle $|z| = R$. This is the approach we take to determine the asymptotic fraction of tautologies and many other classes of formulas among all formulas of a given size.

The main tool used to obtain limits of the fraction of two sequences which are described by generating functions will be the following result, due to Szegő [24] [Thm. 8.4], see as well Wilf [26] [Thm. 5.3.2 page 181]. We can see versions of Szegő lemma in action in papers [30], [31] concerning asymptotic probabilities in logic. The second powerful tool we need is so called Drmota-Lalley-Woods theorem which has been developed independently by Drmota in [5], Lalley in [18] and Woods in [28] to study problems involving the enumeration of families of plane trees or context-free languages and finding their asymptotic behaviour as solutions of positive algebraic systems. The best presentation of Drmota-Lalley-Woods theorem can be found in Flajolet and Sedgewick in [6] [pp. 446-451]. Excellent overview of the method is due to Daniele Gardy in [9] [Chapter 4, pp 15-16].

Theorems and lemmas in the next chapter 5.3 are proved using Szegő lemma. The result mentioned in Theorems 28 and 31 that the limiting ratio $\mu(\text{CL}^k)$ of classical tautologies with k propositional variables exists, requires the use of Drmota-Lalley-Woods theorem.

5.3 Densities of classes of formulas

In this section we wish to summarize results contained in the papers [30] and [7].

Lemma 25. *The asymptotic probability of the fact that a randomly chosen formula is a simple tautology is:*

$$\mu(G^k) = \lim_{n \rightarrow \infty} \frac{G_n^k}{F_n^k} = \frac{4k+1}{(2k+1)^2}$$

Proof. The first proof of this fact can be found in [21]. A simpler one is in Theorem 30 of [30] page 252. \square

Lemma 26. *The density of simple non-tautologies exists and is equal to*

$$\mu(SN^k) = \lim_{n \rightarrow \infty} \frac{SN_n^k}{F_n^k} = \frac{k(k-1)}{(k+1)^2}.$$

For large k , this density is $1 - 3/k + \Theta(1/k^2)$.

Proof. This result was already given in the paper [21, page 586]. The alternative proof can be found in [7] at Proposition 7. \square

Lemma 27. *The density of less simple non-tautologies is equal to*

$$\mu(LN^k) = \lim_{n \rightarrow \infty} \frac{LN_n^k}{F_n^k} = \frac{2k(k-1)^2}{(k+2)^4}.$$

For large k it is equal to $2/k + \Theta(1/k^2)$.

Proof. The long and complicated proof of this fact can be found in chapter 4.3 of [7]. \square

Theorem 28. *Asymptotically (for a large number k of Boolean variables), all classical tautologies are simple and it follows that all intuitionistic tautologies are classical i.e.*

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{\mu(G^k)}{\mu(CL^k)} &= 1, \\ \lim_{k \rightarrow \infty} \frac{\mu^-(INT^k)}{\mu(CL^k)} &= 1. \end{aligned}$$

Proof. We know that for any k , the limiting ratio $\mu(CL^k)$ of classical tautologies with k propositional variables exists. This result is obtained by standard techniques in analysis of algorithms; we skip the details and refer the interested reader to Flajolet and Sedgewick [6] or to Gardy [9]. From the fact presented in Lemma 17

$$G^k \subseteq INT^k \subseteq CL^k \subseteq F^k \setminus (SN^k \cup LN^k)$$

it follows

$$\mu(G^k) \leq \mu^-(INT^k) \leq \mu(CL^k) \leq 1 - \mu(SN^k) - \mu(LN^k).$$

Now our result follows since lower and upper bounds $\mu(G^k)$ and $1 - \mu(SN^k) - \mu(LN^k)$ are equal to $1/k + \Theta(1/k^2)$. \square

Lemma 29. *The density $\mu(PEIRCE^k)$ of Peirce formulas (if it exists) is equal to $\frac{1}{2k^2}$.*

$$\liminf_{n \rightarrow \infty} \frac{PEIRCE_n^k}{F_n^k} = \limsup_{n \rightarrow \infty} \frac{PEIRCE_n^k}{F_n^k} = \frac{1}{2k^2}.$$

Proof. Proof of this fact appears in the paper [12]. \square

Theorem 30. *For the set $EVEN^k$ of formulas we have:*

$$\begin{aligned} \mu^-(EVEN^k) &= \liminf_{n \rightarrow \infty} \frac{EVEN_n^k}{F_n^k} = 0 \\ \mu^+(EVEN^k) &= \limsup_{n \rightarrow \infty} \frac{EVEN_n^k}{F_n^k} = \frac{1}{2^{k-1}}. \end{aligned}$$

Proof. The first equality is trivial since $EVEN_n^k = \frac{c_n}{2^k} \sum_{j=0}^k \binom{k}{j} (k-2j)^n$ is zero by lemma 24 for all odd numbers n . Now suppose that $n = 2m$.

$$\begin{aligned}
\limsup_{2m \rightarrow \infty} \frac{EVEN_{2m}^k}{F_{2m}^k} &= \lim_{2m \rightarrow \infty} \frac{C_{2m}}{2^k} \sum_{j=0}^k \binom{k}{j} \frac{(k-2j)^{2m}}{F_{2m}^k} \\
&= \frac{1}{2^k} \sum_{j=0}^k \lim_{2m \rightarrow \infty} \binom{k}{j} \frac{C_{2m}(k-2j)^{2m}}{k^{2m} C_{2m}} \\
&= \frac{1}{2^k} \lim_{2m \rightarrow \infty} \binom{k}{0} \frac{C_{2m} k^{2m}}{k^{2m} C_{2m}} + \frac{1}{2^k} \lim_{2m \rightarrow \infty} \binom{k}{k} \frac{C_{2m} (-k)^{2m}}{k^{2m} C_{2m}} \\
&= \frac{1}{2^k} + \frac{1}{2^k} = \frac{1}{2^{k-1}},
\end{aligned}$$

where the second and the third line in the above display are equal as

$$\lim_{2m \rightarrow \infty} \binom{k}{j} \frac{C_{2m}(k-2j)^{2m}}{k^{2m} C_{2m}} = 0$$

for every $0 < j < k$. □

The picture below summarizes all theorems involving densities which are needed for proving our next two results on BCK and BCI logics namely Theorems 31 and 32.

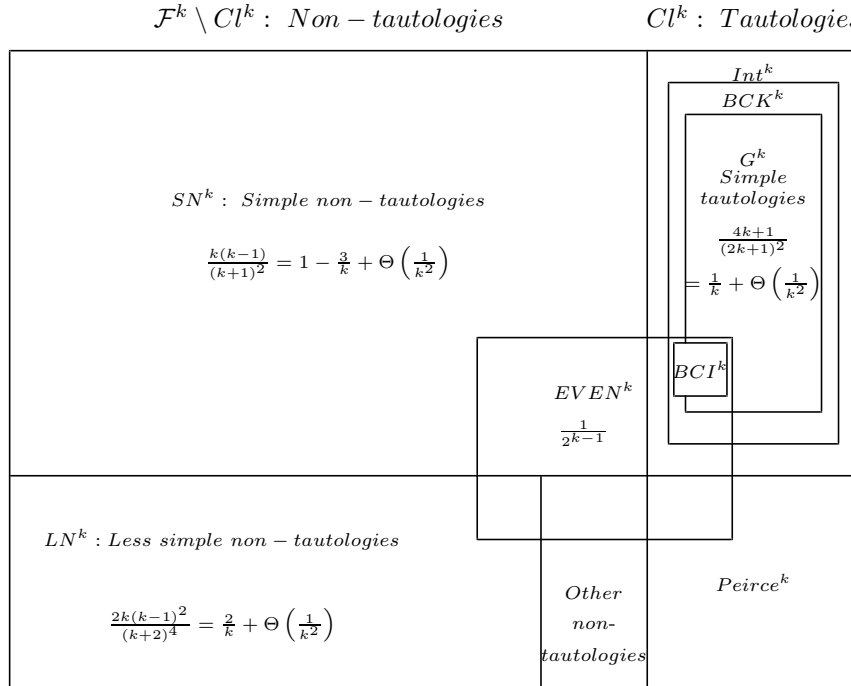


Figure 3: Densities of the sets of formulas

Theorem 31. *Almost every classical tautology is BCK provable i.e.*

$$\lim_{k \rightarrow \infty} \frac{\mu^-(BCK^k)}{\mu(\mathcal{C}l^k)} = 1.$$

Proof. From the fact that $G^k \subseteq BCK^k \subseteq INT^k \subseteq CL^k$, we have

$$\mu(G^k) = \lim_{k \rightarrow \infty} \frac{G_n^k}{F_n^k} \leq \liminf_{n \rightarrow \infty} \frac{BCK_n^k}{F_n^k} \leq \limsup_{n \rightarrow \infty} \frac{INT_n^k}{F_n^k} \leq \lim_{k \rightarrow \infty} \frac{CL_n^k}{F_n^k} = \mu(CL^k).$$

The result follows from the fact that both $\mu(G^k)$ and $\mu(CL^k)$ are equal to $1/k + \Theta(1/k^2)$. \square

Theorem 32. *Almost non BCK provable formula is BCI provable i.e.*

$$\lim_{k \rightarrow \infty} \frac{\mu^+(BCI^k)}{\mu^-(BCK^k)} = 0.$$

Proof. It follows from two facts. Lemma 18 shows $BCI^k \subseteq EVEN^k$ while from Lemma 17 based on Theorem 10 we get $G^k \subset BCK^k$. The rest is based on the calculations from Theorem 30 and Lemma 25.

Therefore:

$$\mu^+(BCI^k) \leq \mu^+(EVEN^k) \leq \frac{1}{2^{k-1}},$$

while

$$\mu^-(BCK^k) \geq \mu(G^k) = \frac{4k+1}{(2k+1)^2}.$$

Finally we have

$$\lim_{k \rightarrow \infty} \frac{\mu^+(BCI^k)}{\mu^-(BCK^k)} \leq \lim_{k \rightarrow \infty} \frac{(2k+1)^2}{2^{k-1}(4k+1)} = 0.$$

\square

Interpretation of result: Weakening rule $\varphi \Rightarrow (\psi \Rightarrow \varphi)$ is much stronger tool to generate formulas than identity rule $\varphi \Rightarrow \varphi$.

Open problems: We do not know if there exist densities $\mu(BCI^k)$ or $\mu(BCK^k)$ of two investigated logics

6 Counting proofs in BCI and BCK logics

In this section we will focus on two special classes BCI and BCK of lambda terms. On the basis of their special structure, we will show how to enumerate BCI and BCK terms of a given size. As the main result the density of BCI terms among BCK terms is computed.

Definition 33. *The size of a lambda term is defined in the following way:*

$$\begin{aligned} \|x\| &= 1 \\ \|\lambda x.M\| &= 1 + \|M\| \\ \|MN\| &= 1 + \|M\| + \|N\|. \end{aligned}$$

As we can see $\|t\|$ is the number of all nodes of lambda tree $G(t)$.

Definition 34. *Let n be an integer. We denote by Λ_n the set of all closed lambda terms up to α conversion of size n . Obviously the set Λ_n is finite. We denote its cardinality by L_n .*

As far as we know, no asymptotic analysis of the sequence L_n has been done. Moreover, typical combinatorial techniques do not seem to apply easily for this task. For the first time the problem of enumerating lambda terms was considered in [25]. In general, counting lambda terms of a given size turns out to be a non-trivial and challenging task. The wide discussion on this problem and some results concerning properties of random terms can be found in [4].

6.1 Enumerating BCI terms

Definition 35. By a_n we denote the number of BCI terms of size n .

Since in BCI terms each lambda binds exactly one variable, in a lambda tree for such a term the number of leaves is equal to the number of unary nodes. In every unary-binary tree the number of leaves is greater by one than the number of binary nodes. Thus the number of BCI terms is positive only if the size is equal to $3k + 2$ (k binary nodes, $k + 1$ unary nodes and $k + 1$ leaves) for $k \in \mathbb{N}$.

Definition 36. By a_n^* we denote the number of BCI terms up to α conversion with n binary nodes.

Obviously, $a_n^* = a_{3n+2}$. Moreover $a_n = 0$ for $n \not\equiv 2 \pmod 3$.

Lemma 37. The sequence (a_n^*) satisfies the recurrence:

$$\begin{aligned} a_0^* &= 1, & a_1^* &= 5, \\ a_n^* &= 6na_{n-1}^* + \sum_{i=1}^{n-2} a_i^* a_{n-i-1}^* \quad , \text{ for } n \geq 2. \end{aligned}$$

Proof. There is only one BCI term of size 2 (no binary nodes): $\lambda x.x$. Moreover there are five terms of size 5 (one binary node): $\lambda xy.xy$, $\lambda xy.yx$, $(\lambda x.x)(\lambda x.x)$, $\lambda x.(\lambda y.y)x$ and $\lambda x.x(\lambda y.y)$. Thus, $a_0^* = 1$ and $a_1^* = 5$.

Let P be a BCI term with $n \geq 2$ binary nodes. Such a term is either in the form of application or in the form of abstraction. Both cases are depicted in Figure 4.

In the first case P is an application of two BCI terms, $P \equiv MN$, where M has i binary nodes and N has $n - i - 1$ binary nodes ($i = 0, \dots, n - 1$). It gives us $\sum_{i=0}^{n-1} a_i^* a_{n-i-1}^*$ possibilities.

In the second case P is in the form of abstraction, $P \equiv \lambda x.M$, and x occurs free in M exactly once. In the tree corresponding to M , the parent of the leaf labeled with x must be a binary node. Thus, we can look at the tree corresponding to M as at the lambda tree for some BCI term Q with $n - 1$ binary nodes with an additional leaf labeled with x . This leaf can be inserted into the tree in two manners, either on the left or on the right. Moreover this can be done in $3n - 1$ ways which is the number of all branches in the tree for Q . Thus, there are $2(3n - 1)a_{n-1}^*$ possibilities of such insertions.

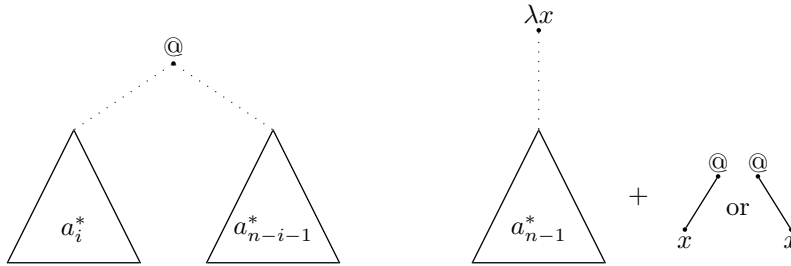


Figure 4: Two ways of obtaining a BCI term with $n \geq 2$ binary nodes

Summing up we get the equation of the lemma. □

Denoting by $A(x)$ the generating function for the sequence (a_n^*) and after basic calculations, we get

$$6x^2 \frac{\partial A(x)}{\partial x} + xA^2(x) + (4x - 1)A(x) + 1 = 0, \quad A(0) = 1.$$

This is a non-linear Riccati differential equation and as such it has a solution which is a non-elementary function.

The sequence (a_n^*) and the function $A(x)$ were studied in [14]. On the basis of that paper we get the asymptotics

$$a_{3n+2} = a_n^* \sim \frac{1}{2\pi} 6^n (n-1)!.$$

First values of (a_n) are the following:

$$0, 0, 1, 0, 0, 5, 0, 0, 60, 0, 0, 1105, 0, 0, 27120, 0, 0, 828250, 0, 0, 30220800, \dots$$

This sequence can be found in the On-Line Encyclopedia of Integer Sequences ([22]) under the number A062980.

6.2 Enumerating BCK terms

Definition 38. Let us denote by b_n the number of BCK lambda terms of size n .

Lemma 39. The sequence (b_n) satisfies the following recursive equation:

$$b_0 = b_1 = 0, \quad b_2 = 1, \quad b_3 = 2, \quad b_4 = 3,$$

$$b_n = b_{n-1} + 2 \sum_{i=0}^{n-3} i b_i + \sum_{i=0}^{n-1} b_i b_{n-i-1} + 1 \quad \text{for } n \geq 5.$$

Proof. There are no terms of size 0 and 1, there is only one BCK term of size 2: $\lambda x.x$, two terms of size 3: $\lambda xy.x$ and $\lambda xy.y$, and three terms of size 4: $\lambda xyz.x$, $\lambda xyz.y$ and $\lambda xyz.z$. Thus $b_0 = b_1 = 0$, $b_2 = 1$, $b_3 = 2$ and $b_4 = 3$.

Let P be a BCK term of size $n \geq 5$. Such a term is either in the form of application or in the form of abstraction where the first lambda binds one variable or in the form of abstraction where the first lambda does not bind any variables.

In the first case P is in the form of application, $P \equiv MN$, where M and N are BCK terms of size, respectively, i and $n - i - 1$ ($i = 0, \dots, n - 1$). It gives us $\sum_{i=0}^{n-1} b_i b_{n-i-1}$ possibilities.

In the second case P is in the form of abstraction, $P \equiv \lambda x.M$ and x occurs free in M exactly once. There are two subcases here: either $M \equiv \lambda x_1 \dots x_{n-2}.x$ or M is a term built of a BCK term Q of size $i = 2, \dots, n - 3$ with an additional term $\lambda x_1 \dots x_{n-i-3}.x$ inserted on one of its branches or on the branch joining λx with Q . Since in Q there are $i - 1$ branches and the additional term can be inserted either on the left or on the right, this case gives us $1 + 2 \sum_{i=2}^{n-3} i b_i$ possibilities. Both subcases are presented in Figure 5.

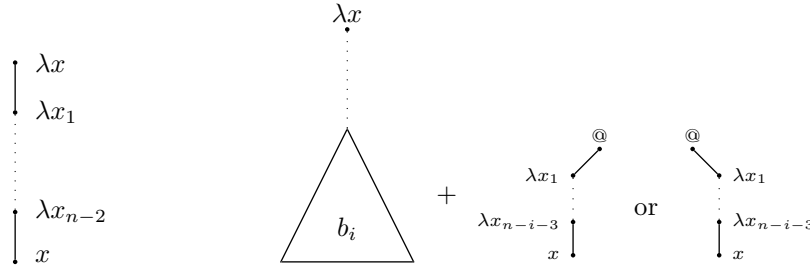


Figure 5: Second case of the construction of a BCK term of size $n \geq 5$

In the third case P is in the form of abstraction, $P \equiv \lambda x.M$ and x does not occur free in M . The number of such terms of size n is equal to the number of all BCI terms of size $n - 1$. Thus, it gives us b_{n-1} possibilities.

Summing up we get the equations of the lemma. \square

Denoting by $B(x)$ the generating function for the sequence (b_n) and after basic calculations, we get

$$2x^4 \frac{\partial B(x)}{\partial x} + (x - x^2)B^2(x) - (1 - x)^2 B(x) + x^2 = 0, \quad B(0) = 0.$$

Again we obtained a non-linear Riccati differential equation. Unfortunately, this time we do not know the solution.

First values of (b_n) are the following:

$$0, 0, 1, 2, 3, 9, 30, 81, 225, 702, 2187, 6561, 19602, 59049, 177633, 532170, 1594323, \dots$$

Also this sequence can be found in [22] under the number A073950.

6.3 Density of BCI in BCK

Our main goal is to compute the density of BCI in BCK. Since there are no BCI terms of size $n \not\equiv 2 \pmod 3$ (which is not true in the case of BCK terms), if the density exists, then it is 0.

Let us observe that each BCK term can be obtained from a BCI term with some additional (possibly none) lambdas. This observation allows us to obtain a formula for b_n depending on the sequence (a_n) .

Lemma 40. *The number of possible ways of choosing k out of the n elements with repetition is equal to $\binom{n+k-1}{n-1}$.*

Lemma 41. *For $k \in \mathbb{N}$ the following formula holds*

$$b_{3k+2} = \sum_{i=0}^k \binom{3k}{3i} a_{3i+2}.$$

Proof. Each BCK term of size $3k+2$ can be obtained from a BCI term of size $3i+2$ ($i = 0, \dots, k$) by inserting $3k+2-3i-2 = 3k-3i$ additional lambdas. There are $3i+1$ branches in the tree for a BCI term of size $3i+2$, thus the number of insertions corresponds to the number of $3k-3i$ -element combinations with repetition of a $3i+1$ -element set and thus, by Lemma 40, it is equal to $\binom{3k}{3k-3i} = \binom{3k}{3i}$. \square

Theorem 42. *The density of BCI terms among BCK terms equals 0.*

Proof. By the asymptotics of (a_{3k+2}) and by Lemma 41, we get

$$\frac{a_{3k+2}}{b_{3k+2}} \leq \frac{a_{3k+2}}{a_{3k+2} + \binom{3k}{3} a_{3k-1}} \xrightarrow{k \rightarrow \infty} 0. \quad \square$$

References

- [1] Blok, W.J., Pigozzi, D. Algebraizable Logics, *Memoirs of the A.M.S.*, 77, Nr. 396, 1989.
- [2] H.B. Curry, R. Hindley and J.P. Seldin, *Combinatory Logic II*, North Holland, 1972.
- [3] B. Chauvin, P. Flajolet, D. Gardy and B. Gittenberger. And/Or trees revisited. *Combinatorics, Probability and Computing*, 13(4-5): 475–497, 2004.

- [4] René David, Katarzyna Grygiel, Jakub Kozik, Christophe Raffalli, Guillaume Theyssier, Marek Zaionc, *Some properties of random λ -terms*, draft published at Computing Research Repository (CoRR), 2009, arXiv:0903.5505v1.
- [5] M. Drmota. Systems of functional equations, *Random Structures and Algorithms*, 10, pp103-124, 1997.
- [6] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.
- [7] H. Fournier, D. Gardy, A. Genitrini and M. Zaionc. *Classical and intuitionistic logic are asymptotically identical*. In Proc. *Internat. Conf. on Computer Science Logic (CSL)*, LNCS Vol. 4646: 177–193, 2007.
- [8] Ghislain R. Franssens, *On a Number Pyramid Related to the Binomial, Deleham, Eulerian, MacMahon and Stirling number triangles*, Journal of Integer Sequences, Volume 9, 2006, Article 06.4.1 pp 1-34.
- [9] D. Gardy. Random Boolean expressions. *Colloquium on Computational Logic and Applications*, Chambéry (France), June 2005. Proceedings in *Discrete Mathematics and Theoretical Computer Science*, AF: 1–36, 2006.
- [10] D. Gardy and A. Woods. And/Or tree probabilities of Boolean function. First Colloquium on the Analysis of Algorithms, Barcelona, 2005. Proceedings in *Discrete Mathematics and Theoretical Computer Science*, AD: 139–146, 2005.
- [11] A. Genitrini and J. Kozik. Quantitative comparison of Intuitionistic and Classical logics – full propositional system. *International Symposium on Logical Foundations of Computer Science*, Florida (USA), Jan. 2009. Proceedings LNCS Vol. 5407: 280–294.
- [12] A. Genitrini, J. Kozik and G. Matecki. On the density and the structure of the Peirce-like formulae. In Proc. *Fifth Colloquium on Mathematics and Computer Science*, Blaubeuren, Germany, September 2008. DMTCS Proceedings.
- [13] J. Roger Hindley, *Basic Simple Type Theory*. Cambridge Tracts in Theoretical Computer Science, Vol. 42, Cambridge: Cambridge University Press, 1997.
- [14] Svante Janson, *The Wiener index of simply generated random trees*. Random Structures Algorithms 22 (2003) 337-358.
- [15] Z. Kostrzycka and M. Zaionc. Statistics of intuitionistic versus classical logic. *Studia Logica*, 76(3): 307–328, 2004.
- [16] T. Kowalski, *Tomasz Self-implications in BCI*. Notre Dame Journal of Formal Logic 49 (2008), no. 3, 295–305
- [17] J. Kozik. Subcritical pattern languages for And/Or trees. In Proc. *Fifth Colloquium on Mathematics and Computer Science*, Blaubeuren, Germany, september 2008. DMTCS Proceedings.
- [18] S.P. Lalley. Finite range random walk on free groups and homogeneous trees, *Ann. Probab.*, 21(4), pp 2087-2130, 1993.
- [19] H. Lefmann and P. Savický. Some typical properties of large And/Or Boolean expressions. *Random Structures and Algorithms*, 10: 337–351, 1997.
- [20] G. Matecki. Asymptotic density for equivalence. *Electronic Notes in Theoretical Computer Science*, 140: 81–91, 2005.

- [21] M. Moczurad, J. Tyszkiewicz and M. Zaionc. Statistical properties of simple types. *Mathematical Structures in Computer Science*, 10(5): 575–594, 2000.
- [22] Neil J. A. Sloane. The on-line encyclopedia of integer sequences. Available at: <http://www.research.att.com/~njas/sequences/index.html>
- [23] R. P. Stanley, *Enumerative Combinatorics*, Wadsworth and Brooks/Cole, 1986.
- [24] G. Szegő, Orthogonal polynomials, fourth ed., AMS, Colloquium Publications, 23, Providence, 1975.
- [25] Jue Wang, *Generating Random Lambda Calculus Terms*. Available at: <http://cspeople.bu.edu/juewang/research.html>
- [26] H. Wilf. *Generatingfunctionology*. Second edition, Academic Press, Boston, 1994.
- [27] A. Woods. On the probability of absolute truth for And/Or formulas. *The Bulletin of Symbolic Logic*, 12(3): 523, 2006.
- [28] A. Woods. Coloring rules for finite trees, and probabilities of monadic second order sentences, *Random Structures and Algorithms*, 10, pp453-485, 1997.
- [29] Wroński, A. BCK-algebras do Not Form a Variety, *Math. Japonica*, 28, pp. 211-213, 1983.
- [30] M. Zaionc. On the asymptotic density of tautologies in logic of implication and negation. *Reports on Mathematical Logic*, 39: 67–87, 2005.
- [31] M. Zaionc. Probability distribution for simple tautologies. *Theoretical Computer Science*, 355(2): 243–260, 2006.