Associative and commutative tree representations for Boolean functions☆

Antoine Genitrini[a], Bernhard Gittenberger[b], Veronika Kraus[b], Cécile Mailler[c]

[a] *Laboratoire LIP6, CNRS UMR 7606 and Université Pierre et Marie Curie, 4 place Jussieu, 75252 Paris cedex 05, France.*
[b] *Technische Universität Wien, Wiedner Hauptstrasse 8-10/104, A-1040 Wien, Austria.*
[c] *Laboratoire de Mathématiques de Versailles, CNRS UMR 8100 and Université de Versailles Saint-Quentin-en-Yvelines, 45 avenue des États-Unis, 78035 Versailles, France.*

Abstract

Since the 1990s, the probability distribution on Boolean functions, induced by some random formulas built upon the connectives *And* and *Or*, has been intensively studied. These formulas rely on plane binary trees. We extend all the results, in particular the relation between the probability and the complexity of a function, to more general formula structures: non-binary or non-plane trees. These formulas satisfy the natural properties of associativity and commutativity.

Keywords: Boolean functions, Probability distribution, Random Boolean formulas, Random trees, Asymptotic ratio, Analytic combinatorics.

1. Introduction

Since the 1980s, several papers have focused on probability distribution on Boolean functions induced by random Boolean formulas. We first mention the result of Valiant [24] who constructs a small formula that with high probability represents the Boolean function `Majority`. The method he developed, often called the probabilistic amplification, has then been adapted to build other Boolean functions [2, 7, 17, 23]. The main goal of such studies was to build explicitly a small formula (of size polynomial in the number of variables) for important Boolean functions. All these results are based on very constrained Boolean formulas: the formulas, seen as trees, are balanced and the labelling of the internal nodes is very regular. Later, some results on larger classes of formulas have been obtained, still based on the approach of amplification: [3, 10, 5].

During the 90s, other authors [21, 19] aimed at defining some "natural" probability distributions for Boolean functions based on large random Boolean formulas seen as trees. In these papers no structural constraints are imposed. The internal nodes of the trees are usually labelled by two connectives *And* and *Or* and the external nodes by symbols taken from fixed set of literals. The support of the resulting probability distribution on Boolean functions is the whole set of functions and no more a distribution concentrated on a small subset of functions, like the one of Valiant.

Other papers appeared during the last 15 years: their central goal was to obtain quantitative results from a logic point of view. The first result in this direction has been obtained in 2000 by [20]. It is based on formulas built with the single connective *Implication* and is dedicated to the study of the quantitative ratio of intuitionistic logic within classical logic. The paper presents exact results for the logics induced by a very small number of variables and states a conjecture

☆This research was partially supported by the A.N.R. project *BOOLE*, 09BLAN0011, and by the P.H.C. Amadeus project *Probabilities and tree representations for Boolean functions*. The Austrian authors' work has been supported by *FWF (Austrian Science Foundation), National Research Area S9600, grant S9604* and *ÖAD, grants F03/2010 and F03/2013*.

Email addresses: `antoine.genitrini@lip6.fr` (Antoine Genitrini), `gittenberger@dmg.tuwien.ac.at` (Bernhard Gittenberger), `veronika.kraus@gmx.at` (Veronika Kraus), `cecile.mailler@math.uvsq.fr` (Cécile Mailler)

on the asymptotic behaviour of the ratios of both logics, when the number of variables tends to infinity. The conjecture has then been proved in [13].

This model, based on a single connective, has then been studied in detail in order to understand the behaviour of the whole probability distribution on Boolean functions. The first results on tautologies [13] have proven to be crucial for the study of the whole distribution. The complete study by Fournier *et al.* [11, 12] has linked together the complexity and the probability of a function.

In parallel, models based on Boolean formulas built with two connectives, *And* and *Or*, have been studied. First, Lefmann and Savický [19] established some bounds for the probability of a function, bounds that are linked to the complexity of the functions. These bounds have been improved by Chauvin *et al.* [4] where other models based on Galton-Watson branching processes have been studied as well. Then Kozik [18] has developed a powerful tool based on pattern languages that allows to classify and count large trees according to some structural constraints. Using this tool he managed to compute the asymptotic order of the probability of a function. Both implicational and *And*/*Or* models exhibit the same relation between complexity and probability and, though the way to prove it is not at all the same, the same paradigm is underlying. Namely, almost all trees computing a fixed function can be constructed in a particular way: Start with a minimal tree and attach a large tree such that the function computed is not changed.

As pointed out by Gardy [14] the results discussed above have a fundamental weakness. All models use plane binary trees as their underlying tree model. This implies that formulas which should be considered the same are counted separately in the models: Indeed, since *And* and *Or* are commutative and associative operations, the underlying trees should neither be plane nor binary. Similarly, plane trees are not appropriate for the implication model since the premises of an implicational formula can be interchanged without changing the function. This issue was addressed in [16] where a model of *Implication* which is insensitive to the commutation of premises has been studied.

This paper aims at a thorough analysis of the relation between complexity and probability of a Boolean function given by a large random *And*/*Or*-formula as well as at the study of the influence of associativity and commutativity on the behaviour of the model. Thus we will present results for four models: Formulas with or without associativity and with or without commutativity of the connectives. We will derive precise asymptotic results (including numerical constants) for the probability of functions of smallest complexity (literals and constants) as well as the asymptotic behaviour for functions of higher complexity. The paradigm mentioned above (a typical tree is a minimal tree expanded once) still holds for all our models. In this paper we also analyse where such expansions can take place which enables us to derive bounds for the multiplicative constants of the asymptotic expressions. Our method would allow also the precise computation of the constants in this case, though the derivation would be much more involved. The analysis will utilize and extend Kozik's theory of pattern languages [18]. This method was designed for and successfully applied to the binary plane case. However, the non-binary cases require a modification of the method and in the non-plane case there are no exact formulas available any more, but only approximate ones. We have to utilize Pólya's enumeration theory which makes the analysis of the models technically more difficult. Moreover, we have to work with more general pattern languages and introduce semi-planar structures (which we call mobiles) in Sections 3.3 and 3.4. Unfortunately, these pattern languages are not subcritical any more which was a crucial property in the analysis of the binary plane case. For a global reference on non-plane tree-structures and the techniques that are necessary the reader can take Drmota's book [6].

The results for the first of these models (neither associative nor commutative) are partially known ([25, 18]). However, for comparison and in order to put all the models under a common roof, we will include this model as well.

The paper is organized as follows. Section 2 is dedicated to introduce the whole context of Boolean formulas seen as trees and presents the models and probability distributions we will study. Then the complete study of the distributions is presented. It is decomposed in three sections: Sections 3, 4 and 5. Each section is presented is the same way. First we present an overview of the corresponding result of Kozik in the case of binary and plane formulas in order to point out precisely the technical arguments that must be adapted to address our context of non-binary or non-plane formulas. And then we prove the generalised versions of the key-tools we need. We will prove that the probability of a given Boolean function is asymptotically proportional to a power of the number of allowed variables with exponent related to the complexity. The results are stated in Section 5, Theorems 5.3, 5.8 and 5.9. Moreover, we derive narrow bounds for proportionality factors for the probability of any Boolean function and the proofs of these theorems exhibit what most of the formulas for a fixed function look like.

2. Associative and commutative trees: definitions, generating functions

Kozik [18] has shown that in binary plane trees the order of magnitude of the limiting probability of a given Boolean function is related to its complexity. We generalise this result and therefore define the complexity of a function by the following:

Definition 2.1. An *And/Or tree* is a labelled tree, where each internal node is labelled with one of the connectors $\{\wedge, \vee\}$ and each leaf with one element of the set of literals $\mathcal{L}_n = \{x_1, \bar{x}_1 \dots, x_n, \bar{x}_n\}$. We define the *size* of an And/Or tree to be its number of leaves. The set of *variables* is $\mathcal{X}_n := \{x_1, \dots, x_n\}$.

Definition 2.2. The *complexity* $L(f)$ of a non-constant function f (i.e. $f \notin \{True, False\}$) is given by the size of a smallest And/Or tree computing f (in the rest of the paper such trees will be called *minimal* for f), while we define the complexity of $True$ and $False$ to be $L(True) = L(False) = 0$.

As it will be clear later, the complexity of a function does not depend on the chosen tree model.

Definition 2.3. We are considering a set $\mathcal{T}_{m,n}$ of And/Or trees of size m. Let $\mathbb{U}_{m,n}$ be the uniform distribution on $\mathcal{T}_{m,n}$, and $\mathbb{P}_{m,n}$ its image on the set of Boolean functions. We call

$$\mathbb{P}_n = \lim_{m\to\infty} \mathbb{P}_{m,n}$$

the limiting distribution.

Remark: In all models we will take into consideration, the probability of a function f is equal to the one of its negation. In fact, a tree computing f can be relabelled in the following way: each connector is substituted by the other one and each literal by its negation ($x \to \bar{x}$ and $\bar{x} \to x$). The new tree we obtain belongs to the same model as t and computes the function $\bar{f}$.

At first, we will present the result proven by Kozik. This result will be generalised in the subsequent sections of the paper.

2.1. The classical model.

First, let us consider the set $\mathcal{T}$ of binary plane trees, whose internal nodes are labelled with $\wedge$ or $\vee$, and whose external nodes are labelled with literals chosen in $\mathcal{L}_n$: each such tree computes a Boolean function on n variables. We denote by $T(z) = \sum_{m\geq 0} T_m z^m$ the generating function enumerating this set of trees[1], and by $T_f(z)$ the generating function of such trees computing the Boolean function f. Let us remind some well known results about this generating function. These rely on the so-called symbolic method (see [9, Ch. I] for an introduction):

[1] More generally, in the rest of this paper, a generating function and its coefficients will be denoted by the same capital letter $Z(z)$ for the generating function and Z_m for its coefficients.

Proposition 2.4. *Let $\mathcal{T}$ denote the set of binary* And/Or *trees and $\mathcal{Z} = \mathcal{L}_n$ the set of leaves. Then a tree is either a single leaf or an* And*-node with two binary trees attached to it or an* Or*-node with two binary trees attached to it. This gives rise to the symbolic equation*

$$\mathcal{T} = \mathcal{Z} \mid \mathcal{T} \wedge \mathcal{T} \mid \mathcal{T} \vee \mathcal{T}. \tag{1}$$

Thus the generating function $T(z)$ verifies $T(z) = 2nz + 2T(z)^2$ and therefore we have

$$T(z) = \frac{1 - \sqrt{1 - 16nz}}{4}$$

and the singularity ρ_n of $T(z)$ is $\frac{1}{16n}$.

Let us consider the uniform distribution on the set of trees of size m and then the probability distribution $\mathbb{P}_{m,n}$ it induces on the set $\mathcal{F}_n$ of Boolean functions on n variables. The limit of this distribution when m tends to infinity (cf. Definition 2.3), denoted by $\mathbb{P}_n$ has already been studied, in particular by Lefmann and Savický [19], Chauvin *et al.* [4] and Kozik [18], who has shown the following theorem.

Theorem 2.5 (Kozik [18])**.** *Let f be a Boolean function. Then,*

$$\mathbb{P}_n(f) \sim \frac{\lambda_f}{n^{L(f)+1}} \text{ as } n \to \infty,$$

where $L(f)$ is the complexity of f, i.e. the size of a minimal tree computing f, and λ_f is a constant depending on f, which will be specified later in this paper.

Remark: Note that in [18], the result $\mathbb{P}_n(f) = \Theta\left(\frac{1}{n^{L(f)+1}}\right)$ is rigorously proven for the binary plane model, while the actual existence of the constant is suggested. In Section 3 we will compute λ_f for the constant functions $True$ and $False$ (already shown in [25] and later in [18] as well), in Section 4 for literals. We show its existence in general and derive bounds in Section 5.

Definition 2.6. A variable x is essential for a function $f = f(x, x_1, \dots, x_{n-1})$ if there exists an assignment of $True$ or $False$ to the variables $x_1, \dots, x_{n-1}$, which we denote by $\underline{x}_0$, such that $f(True, \underline{x}_0) \neq f(False, \underline{x}_0)$.

Remark: An essential variable of f appears in every tree representation of f.

Remark: Note that in this theorem, f (and thus $L(f)$) is fixed, and n tends to infinity. The set of essential variables of the function is finite (and does not depend on n).

First of all, let us define associative trees, commutative trees and then associative and commutative trees, and the induced distributions on the set of Boolean functions $\mathcal{F}_n$.

2.2. The associative plane model.

Definition 2.7. An *associative* tree is a plane tree where each node has out-degree chosen in $\mathbb{N} \setminus \{1\}$. A *labelled associative* tree is an associative tree in which each external node has a label in $\mathcal{L}_n$ and each internal node has an $\wedge$-label or an $\vee$-label but cannot have the same label as its father. We denote by $\mathcal{A}$ the family of associative trees and by $\mathcal{A}_m$ the set of such trees of size m.

Hence these trees are *stratified*: the root can be labelled either by $\wedge$ or $\vee$ and it determines the labels of all other internal nodes.

We denote by $\mathbb{P}^a_n = \lim_{m\to\infty} \mathbb{P}^a_{m,n}$ the limiting distribution of Boolean functions induced by associative And/Or trees. Our aim is to compare the limiting distributions $\mathbb{P}^a_n$ and $\mathbb{P}_n$.

The generating function enumerating associative trees is given by $A(z) = \hat{A}(z) + \check{A}(z) - 2nz$, where $\hat{A}$ (resp. $\check{A}$) is the generating function of associative trees rooted at an $\wedge$-node (resp. an $\vee$-node) or is a single leaf. Note that $\hat{A}(z) = \check{A}(z)$ and,

$$\hat{A}(z) = 2nz + \sum_{k\geq 2} \check{A}(z)^k = 2nz + \frac{\hat{A}^2(z)}{1 - \hat{A}(z)}.$$

Therefore,

$$A(z) = \frac{1}{2}\left(1 - 2nz - \sqrt{1 - 12nz + 4n^2z^2}\right) \tag{2}$$

and its dominant singularity is

$$\alpha_n = \frac{3 - 2\sqrt{2}}{2n}.$$

Moreover, $A(\alpha_n) = \sqrt{2} - 1$.

Remark: Thanks to the Drmota-Lalley-Woods theorem (well presented in [9, Chapter 8]), we can show that $P^a_{m,n}$ has indeed a limit when m tends to infinity. We denote by $\hat{A}_f(z)$ (resp. $\check{A}_f(z)$) the generating function enumerating associative trees computing f, whose roots are labelled by $\wedge$ (resp. $\vee$) or a literal. These generating functions satisfy the following system:

$$\begin{cases} \hat{A}_f(z) = z\mathbb{1}_{\{f \text{ lit.}\}} + \displaystyle\sum_{i=2}^{\infty} \sum_{\substack{g_1,\dots,g_i,\\ g_1\wedge\cdots\wedge g_i = f}} \check{A}_{g_1}(z)\cdots\check{A}_{g_i}(z) \\ \check{A}_f(z) = z\mathbb{1}_{\{f \text{ lit.}\}} + \displaystyle\sum_{i=2}^{\infty} \sum_{\substack{g_1,\dots,g_i,\\ g_1\vee\cdots\vee g_i = f}} \hat{A}_{g_1}(z)\cdots\hat{A}_{g_i}(z). \end{cases}$$

The Drmota-Lalley-Woods theorem says, roughly speaking, that generating functions satisfying a system of functional equation have a dominant singularity of the same type. By transfer theorems (see [8]) this implies similar behaviour of their coefficients and eventually the existence of the limiting distribution (cf. Definition 2.3). For a similar system of functional equations it was shown in [12, Section 3] that all assumptions of the Drmota-Lalley-Woods theorem indeed hold.

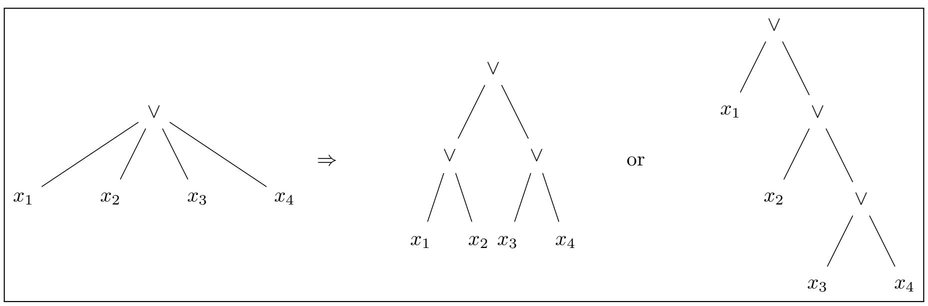


Figure 1: Two of the five possible binary trees obtained from the associative tree.

2.3. The commutative binary model.

Definition 2.8. A *labelled commutative* tree on n variables is a non-plane binary tree where every internal node is labelled with one of the connectors $\{\wedge, \vee\}$ and every leaf is labelled by a literal from $\mathcal{L}_n = \{x_i, \bar{x}_i, i = 1, \dots, n\}$. We denote this family of trees by $\mathcal{C}$.

We consider the distribution $\mathbb{P}^c_{m,n}$ induced over the set of Boolean functions of n variables by the uniform distribution over such trees of size m.

Binary commutative trees fulfill the same symbolic equation as in the plane case (c.f. (1)) but because of commutativity, the generating function of all commutative trees on n variables, counting leaves, is given implicitly by

$$C(z) = 2nz + C(z)^2 + C(z^2), \tag{3}$$

where the term $\frac{1}{2}(C(z)^2 + C(z^2))$ tracks a possible symmetry if both subtrees of the root are identical. See Gardy [14] for details on this model of expressions and Pólya and Read [22] for more general ideas. The system of equations for the generating functions $C_f(z)$ computing a given Boolean function f is given by

$$C_f(z) = z\mathbb{1}_{\{f \text{ lit.}\}} + \frac{1}{2}\sum_{\substack{g,h\neq f\\ g\wedge h = f}} C_g(z)C_h(z) + \frac{1}{2}\sum_{\substack{g,h\neq f\\ g\vee h = f}} C_g(z)C_h(z) + C_f(z)^2 + C_f(z^2).$$

We can prove all assumptions of the Drmota-Lalley-Woods theorem, hence we conclude that all the $(C_f(z))$ and $C(z)$ have the same singularity γ_n, and therefore $\mathbb{P}^c_{m,n}$ converges to a limiting probability distribution P^c_n, when m tends to infinity.

2.4. The commutative associative model.

Definition 2.9. Finally we define *general labelled* trees as commutative and associative trees, with internal nodes labelled by $\wedge$ or $\vee$ (with the condition that father and sons cannot have the same label), and external nodes labelled by literals chosen in $\mathcal{L}_n$. We denote by $\mathcal{P}$ this family of trees.

As in the other models, we consider the distribution $\mathbb{P}^{a,c}_{m,n}$ induced over the set of Boolean functions by the uniform distribution over such trees of size m.

Let $P(z) = \sum_m P_m z^m$ be the generating function of general trees, and $\hat{P}(z)$ (resp. $\check{P}(z)$) the generating function of general trees rooted by $\wedge$ (or by $\vee$, resp.) or are a leaf. We have

$$P(z) = \hat{P}(z) + \check{P}(z) - 2nz, \tag{4}$$

with

$$\begin{cases} \hat{P}(z) = \exp\left(\sum_{i\geq 1}\frac{\check{P}(z^i)}{i}\right) - 1 - \check{P}(z) + 2nz \\ \check{P}(z) = \exp\left(\sum_{i\geq 1}\frac{\hat{P}(z^i)}{i}\right) - 1 - \hat{P}(z) + 2nz. \end{cases} \tag{5}$$

Moreover, the generating functions $\hat{P}_f(z)$ and $\check{P}_f(z)$ of general trees computing f satisfy the following system:

$$\begin{cases} \hat{P}_f(z) = z\mathbb{1}_{\{f \text{ lit.}\}} + \sum_{l=2}^{\infty}\sum_{\substack{g_1,\dots g_i,\\ g_1\wedge\ldots\wedge g_l = f}}\prod_{j=1}^{l}\left(\exp\left(\sum_{i\geq 1}\frac{\check{P}_{g_j}(z^i)}{i}\right) - 1\right) \\ \check{P}_f(z) = z\mathbb{1}_{\{f \text{ lit.}\}} + \sum_{l=2}^{\infty}\sum_{\substack{g_1,\dots g_i,\\ g_1\wedge\ldots\wedge g_l = f}}\prod_{j=1}^{l}\left(\exp\left(\sum_{i\geq 1}\frac{\hat{P}_{g_j}(z^i)}{i}\right) - 1\right). \end{cases}$$

Thus, we can check the hypothesis of the Drmota-Lalley-Woods theorem and conclude that the limiting distribution $\mathbb{P}^{a,c}_n$ of $\mathbb{P}^{a,c}_{m,n}$, when m tends to infinity, exists, and moreover, that all the $\hat{P}_f, \check{P}_f$, $\hat{P}$ and $\check{P}$ have the same singularity, denoted by δ_n.

In the next sections of the paper, we will show that Theorem 2.5 still holds in the associative or commutative cases.

First, we show in Section 3 that the limiting ratio of tautologies is of order $\frac{1}{n}$, we compute explicitly the limit of $\mathbb{P}_n(True)$ when n tends to infinity for the different models. If these limits were the same, we could not conclude anything, but in fact they are all different, which permits us to conclude that asymptotically, when n tends to infinity, the probability distributions induced by the various models are all different. In Section 4, we extend our results to the limiting probabilities of functions which are literals. In all models, the asymptotic ratio is of order $\frac{1}{n^2}$ when n tends to infinity, but the limiting ratios are different from one model to the other. Finally, we generalise Theorem 2.5 in Section 5.

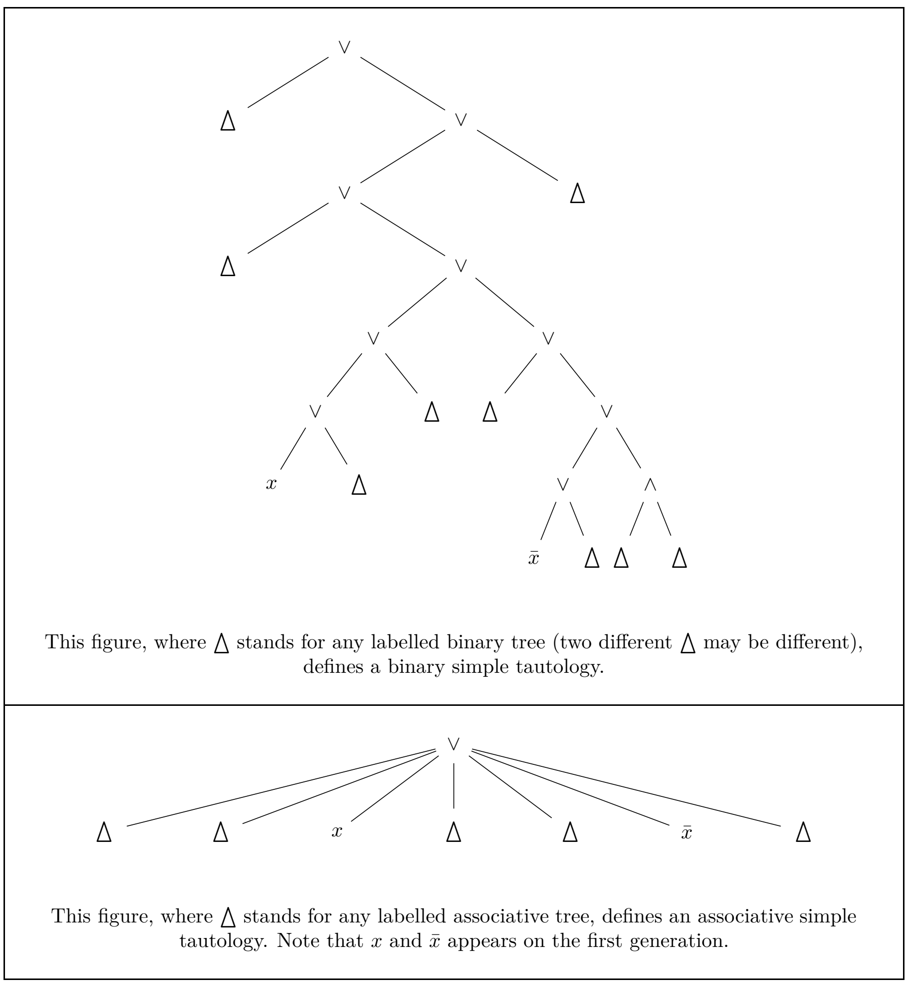
3. Limiting probability of tautologies

In this section we compute the limiting probability of the constant function $True$. We recall that trees computing the function $True$ are called *tautologies*.

Definition 3.1. In a tree, if the path from the root to a leaf crosses only $\vee$- nodes, then this path will be called an $\vee$-*only-path*. We extend the definition to the case such that the leaf is equal to the root (i.e. the tree has size 1).

As suggested by Kozik's results, the limiting probability of tautologies reduces to the limiting probability of so-called *simple tautologies*, defined by the following:

Definition 3.2. A *simple tautology realised by* x_i, $i = 1 \dots n$, is a Boolean expression which has the shape $x_i \vee \bar{x}_i \vee f$ for some Boolean function f, i.e. there exists a leaf labelled by x_i and a leaf labelled by $\bar{x}_i$, both connected to the root by an $\vee$-only-path (c.f. Figure 2). A *simple tautology* is a simple tautology realised by any literal $x \in \{x_1, \dots, x_n\}$. We denote by ST_m the number of simple tautologies of size m (on n variables, n is omitted for simplicity), and $ST = \cup_m ST_m$.



This figure, where Δ stands for any labelled binary tree (two different Δ may be different), defines a binary simple tautology.

This figure, where Δ stands for any labelled associative tree, defines an associative simple tautology. Note that x and $\bar{x}$ appears on the first generation.

Figure 2: Simple tautologies.

Definition 3.3. Let $\mathcal{V}$ be a set of variables and $ST_m(\mathcal{V})$ be the set of simple tautologies realised by every $x \in \mathcal{V}$ but not by any other variable $y \notin \mathcal{V}$.

- $K_{1,m}$ is the set of simple tautologies that are realised by exactly one variable:

$$K_{1,m} = \biguplus_{i=1}^{n} ST_m(\{x_i\}),$$

- $K_{2,m}$ is the set of simple tautologies that are realised by exactly two different variables:

$$K_{2,m} = \biguplus_{\substack{i,j=1 \\ i \neq j}}^{n} ST_m(\{x_i, x_j\}),$$

$\vdots$

- $K_{n,m}$ is the set of simple tautologies that are realised by exactly n different variables:

$$K_{n,m} = ST(\{x_1, \dots, x_n\}).$$

Let $ST^x(z)$ denote the generating function of simple tautologies realised by x and $G(z) = ST^{x_1}(z)+ST^{x_2}(z)+\cdots+ST^{x_n}(z) = nST^x(z)$. Obviously, $\forall m \in \mathbb{N}$, $\#K_{1,m} \leq ST_m \leq G_m$, because some tautologies are counted several times in G. We get $G_m = \#K_{1,m}+2\cdot\#K_{2,m}+\cdots+n\cdot\#K_{n,m}$.

To calculate limiting probabilities, we use the singular expansions of the considered generating functions around their dominant singularities.

Lemma 3.4. *Consider the generating function $T(z)$ of a given family $\mathcal{T}$ of And/Or trees together with the generating function $S(z)$ of a subset $\mathcal{S} \subseteq \mathcal{T}$ of such trees. We assume that $T(z)$ and $S(z)$ have the same dominant singularity ρ and a square root singular expansion*

$$T(z) = a_T - b_T\sqrt{1-\frac{z}{\rho}} + \mathcal{O}\left(1-\frac{z}{\rho}\right), \quad S(z) = a_S - b_S\sqrt{1-\frac{z}{\rho}} + \mathcal{O}\left(1-\frac{z}{\rho}\right),$$

around ρ. Then,

$$\lim_{m\to\infty} \frac{S_m}{T_m} = \lim_{z\to\rho} \frac{S'(z)}{T'(z)}.$$

We call this number the limiting ratio *of the set $\mathcal{S}$ counted by $S(z)$.*

Proof. If m tends to infinity, transfer lemmas (c.f. [9]) give

$$\left.\begin{array}{l} S_m \sim \dfrac{b_S}{\sqrt{\pi}} n^{-\frac{3}{2}}\rho^{-m} \\ T_m \sim \dfrac{b_T}{\sqrt{\pi})} n^{-\frac{3}{2}}\rho^{-m} \end{array}\right\} \Rightarrow \frac{S_m}{T_m} \sim \frac{b_S}{b_T}.$$

Derivation of the singular expansions gives

$$S'(z) \sim \frac{b_S}{2}\left(1-\frac{z}{\rho}\right)^{-\frac{1}{2}}, \qquad T'(z) \sim \frac{b_T}{2}\left(1-\frac{z}{\rho}\right)^{-\frac{1}{2}}.$$

Hence the result follows. ☐

Remark: If $\mathcal{S}$ is the set of trees computing a given function f, then the limiting probability of f is equal to the limiting ratio of $\mathcal{S}$ because for all $m \geq 1$,

$$\mathbb{P}_{m,n}(f) = \frac{\#\text{ trees of size } m \text{ computing } f}{\#\text{ all trees of size } m} = \frac{S_m}{T_m}.$$

3.1. Binary plane trees

In the binary plane model, Kozik has shown that asymptotically, when n tends to infinity, all tautologies are simple tautologies. Therefore, to estimate the probability that a binary plane tree computes the function $True$, it suffices to count simple tautologies, and furthermore, thanks to the following proposition, simple tautologies that are realised by only one variable (i.e. the set $K_{1,m}$).

Proposition 3.5. *If n tends to infinity, then*

$$\lim_{m\to\infty} \frac{1}{T_m} \sum_{k=1}^{n} k \cdot \#K_{k,m} = \lim_{m\to\infty} \frac{\#K_{1,m}}{T_m} + \mathcal{O}\left(\frac{1}{n^2}\right).$$

The proof of the proposition is deferred to the end of this section since further technical concepts are required.

Theorem 3.6. *The limiting ratio of simple tautologies, and thus the limiting ratio of tautologies in the binary plane model is*

$$\lim_{m\to\infty} \mathbb{P}_{m,n}(True) = \lim_{m\to\infty} \frac{ST_m}{T_m}\left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right) = \frac{3}{4n} + \mathcal{O}\left(\frac{1}{n^2}\right), \text{ when } n \text{ tends to infinity,}$$

where T_m is the total number of plane binary trees and ST_m is the number of simple tautologies of size m labelled with n variables.

Proof. Let us compute the generating function of simple tautologies. First, let g_x be the generating function of trees containing a leaf labelled by x which is connected to the root by an $\vee$-only-path (c.f. Figure 3) and $\bar{g}_x(z)$ the generating function of trees which are not of such shape. Hence $g_x = T - \bar{g}_x$.

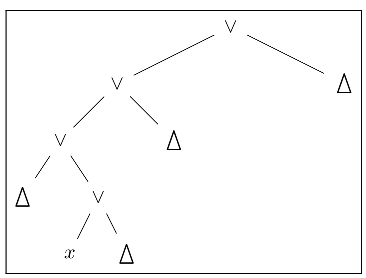


Figure 3: A tree counted by the generating function g_x.

The function $\bar{g}_x$ is given by:

$$\bar{g}_x(z) = T(z)^2 + \bar{g}_x(z)^2 + (2n-1)z.$$

This equation is obtained by decomposing the tree at its root: if the root is labelled by an $\wedge$, the tree is not of the shape depicted in Figure 3 and both subtrees are arbitrary trees. If the root is labelled by an $\vee$, neither of the two subtrees may have the shape of Figure 3. If the root is a single leaf, it must not be labelled by x. By the symbolic method [9] The three cases translate to the three terms in the equation. Solving this equation, using the explicit expression of $T(z)$ given by Proposition 2.4, we get

$$\bar{g}_x(z) = \frac{1}{2} - \frac{\sqrt{2 + 2\sqrt{1-16nz} - 16nz + 16z}}{4},$$

and thus

$$g_x(z) = \frac{\sqrt{2 + 2\sqrt{1-16n} - 16nz + 16z} - \sqrt{1-16nz} - 1}{4}. \qquad (6)$$

Let h_x be the generating function of trees given by $t_1 \vee t_2$ (or $t_2 \vee t_1$), where t_1 is a tree counted by g_x and not by $g_{\bar{x}}$ (therefore it is not a simple tautology) and t_2 is a tree counted by $g_{\bar{x}}$ but not by g_x, i.e. simple tautologies realised by x, where x and $\bar{x}$ must lie in different subtrees of the root (c.f. Figure 4).

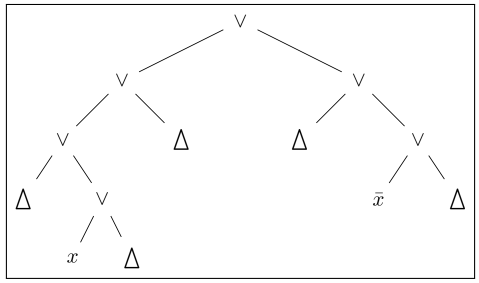


Figure 4: A tree counted by h_x.

Obviously, $h_x(z) = 2(g_x(z) - ST^x(z))^2$. Recall that $ST^x(z)$ is the generating function of simple tautologies realised by the variable x, and $\overline{ST}^x(z)$ be the generating function of trees that are not simple tautologies realised by x. Again by decomposing and analysing the label of the root, we get

$$\overline{ST}^x = T(z)^2 + (\overline{ST}^x(z)^2 - h_x(z)) + 2nz.$$

In particular, if the root is labelled by an $\vee$, neither of the two subtrees can be a simple tautology realised by x and additionally the whole tree cannot be of the shape depicted in Figure 4. Solving this equation, we obtain an explicit expression for $\overline{ST}^x$, and $ST^x(z) = T(z) - \overline{ST}^x(z)$ yields an expression for $ST^x(z)$, where Z denotes $Z := \sqrt{1-16nz}$:

$$ST^x(z) = \frac{1}{4}\left(-1 - Z + 2\sqrt{2 + 2Z - 16nz + 16z} - \sqrt{2 + 2Z - 16nz + 32z}\right).$$

By Proposition 3.5, $\lim_{m\to\infty} \frac{ST_m}{T_m} = \lim_{m\to\infty} \frac{G_m}{T_m} + \mathcal{O}\left(\frac{1}{n^2}\right)$, when n tends to infinity. Due to Lemma 3.4 we can compute the ratio

$$\lim_{m\to\infty} \frac{G_m}{T_m} = \lim_{z\to\frac{1}{16n}} \frac{G'(z)}{T'(z)} = \frac{3}{4n} + \mathcal{O}\left(\frac{1}{n^2}\right)$$

where $G(z) = nST^x(z)$ is given just after Definition 3.3. Thus,

$$\lim_{m\to\infty} \frac{ST_m}{T_m} = \frac{3}{4n} + \mathcal{O}\left(\frac{1}{n^2}\right).$$

Since, when n tends to infinity, asymptotically almost every tautology is a simple tautology, this implies

$$\lim_{m\to\infty} \mathbb{P}_{m,n}(True) = \lim_{m\to\infty} \frac{ST_m}{T_m} + \mathcal{O}\left(\frac{1}{n^2}\right).$$ □

We now go back to Proposition 3.5. In the following, we define pattern languages and some related vocabulary, which can be found in Kozik's paper [18] for the binary case. Interpreting a given And/Or tree as an element from a pattern language, which is possible if pattern and trees have a similar structure, will lead us to the proof of Proposition 3.5.

Definition 3.7. A *pattern language* $\tilde{L}$ is a set of plane trees with internal nodes labelled by $\wedge$ or $\vee$, and external nodes labelled by $\bullet$ or $\Box$. The leaves labelled by $\Box$ are called *placeholders* and those labelled by $\bullet$ are called *pattern leaves*. We define $s(x, y)$ as the generating function of $\tilde{L}$, with x marking the pattern leaves and y marking the placeholders.

Given a pattern language $\tilde{L}$, we will denote by L the set of plane labelled trees with internal nodes labelled by $\wedge$ or $\vee$, and external nodes labelled by literals or placeholders such that, if we replace every literal by a $\bullet$, we obtain a tree of $\tilde{L}$. Therefore, $s(2nx, y)$ is the generating function of L. We call L the *labelled pattern language* associated to $\tilde{L}$.

Given a set of trees $\mathcal{T}$, we define $\tilde{L}[\mathcal{T}]$ (resp. $L[\mathcal{T}]$) as the set of trees obtained by taking an element of $\tilde{L}$ (resp. L) and plugging an element of $\mathcal{T}$ in each placeholder.

Given two pattern languages L and M, we define the *composition* $L[M]$ of L and M by the pattern language obtained by plugging M-patterns into the placeholders of the structures of L. The pattern leaves of $L[M]$ are then both the pattern leaves of L and those of M.

Remark: In order not to overload notation, we will sometimes use the same symbol L for L and $\tilde{L}$ if no confusion arises. Then we will stress verbally whether the labelled or unlabelled version of the pattern language is meant.

Definition 3.8. A (labelled or unlabelled) pattern language L is *unambiguous* if for every family $\mathcal{T}$ every element of $L[\mathcal{T}]$ can be constructed in only one way.

A (labelled or unlabelled) pattern language L with generating function $s(x, y)$ is *subcritical* for $\mathcal{T}$ if the generating function $t(z)$ of $\mathcal{T}$ has a square root singularity ρ and if $s(x, y)$ is analytic in some set $\{(x, y) : |x| \leq \rho + \epsilon,\ |y| \leq t(\rho) + \epsilon\}$.

Definition 3.9. If t is an element of $L[\mathcal{T}]$, we say that t has q *L-repetitions* if q equals the difference between the number of its L-pattern leaves and the number of distinct variables (not literals!) that appear in its L-pattern leaves.

Further, if $\mathcal{V}$ is a fixed subset of the set of variables $\mathcal{X}_n$, we say that t has q *$(L, \mathcal{V})$-restrictions* if q equals the number of its L-repetitions plus the number of variables from $\mathcal{V}$ that appear at least once in the L-pattern leaves of t

Remark: In the typical context needed, $\mathcal{V}$ will be the set of essential variables of the function represented by the tree t. Thus we will call the variables in $\mathcal{V}$ essential variables of t.

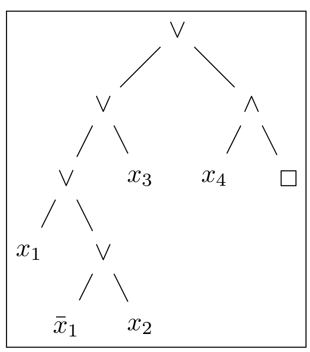


Figure 5: A binary tree with one repetition. Since the tree represents a tautology, none of the variables x_1, x_2, x_3, x_4 is essential for the function. It is therefore natural to set $\mathcal{V} = \emptyset$. The tree has one $(L, \emptyset)$-restriction.

Theorem 3.10. *[18] Let L be a binary unambiguous language which is subcritical for $\mathcal{T}$ and $\mathcal{V}$ be a fixed subset of $\mathcal{X}_n$. We denote by $L[\mathcal{T}]^{[k]}_{m,n}$ (resp by $L[\mathcal{T}]^{[\geq k]}_{m,n}$) the number of elements of $L[\mathcal{T}]$ of size m which have k (resp. at least k) $(L, \mathcal{V})$-restrictions, and by $L[\mathcal{T}]_m$ the number of elements of $L[\mathcal{T}]$ of size m. Then,*

$$\lim_{m\to\infty} \frac{L[\mathcal{T}]^{[\geq k]}_{m,n}}{T_m} \sim \lim_{m\to\infty} \frac{L[\mathcal{T}]^{[k]}_{m,n}}{T_m} \sim \frac{d}{n^k},$$

when n tends to infinity, and d is a constant.

Due to this theorem, we can now prove Proposition 3.5.

Proof of Proposition 3.5. Let us consider the pattern language $S = \bullet|S \vee S| \square \wedge \square$ (c.f. [18]). The set of all trees computing $True$ with exactly i S-restrictions includes K_i. Therefore, thanks to Theorem 3.10, we get

$$\lim_{m\to\infty} \frac{\#K_{i,m}}{T_m} = \mathcal{O}\left(\frac{1}{n^i}\right),$$

when n tends to infinity. Therefore,

$$\lim_{m\to\infty} \frac{2\cdot\#K_{2,m} + \ldots + n\cdot\#K_{n,m}}{T_m} = \mathcal{O}\left(\frac{1}{n^2}\right) + (n-2)\,\mathcal{O}\left(\frac{1}{n^3}\right) = \mathcal{O}\left(\frac{1}{n^2}\right). \qquad \square$$

3.2. Associative plane trees

To compute the limit of $\mathbb{P}^a_n(True)$ when n tends to infinity, we define simple tautologies, and prove that asymptotically every tautology is a simple tautology. Therefore, we will generalise Theorem 3.10 to associative trees.

Theorem 3.11. *The limiting probability of the function $True$ in the associative model is given by*

$$\mathbb{P}^a_n(True) = \lim_{m\to\infty} \mathbb{P}^a_{m,n}(True) = \frac{51-36\sqrt{2}}{n} + \mathcal{O}\left(\frac{1}{n^2}\right).$$

Let us first show that Theorem 3.10 can be generalised to the associative case, and then use it to show Theorem 3.11.

3.2.1. Generalisation of Kozik's theorem to associative trees

Theorem 3.12. *Let $\mathcal{V} \subseteq \mathcal{X}_n$ be a fixed set and L be an unambiguous pattern language with out-degree different from 1, which is subcritical for $\mathcal{A}$. We denote by $L[\mathcal{A}]^{[k]}_{m,n}$ (resp by $L[\mathcal{A}]^{[\geq k]}_{m,n}$) the number of elements of $L[\mathcal{A}]$ of size m which have k (resp. at least k) $(L,\mathcal{V})$-restrictions, and by $L[\mathcal{A}]_m$ the number of elements of $L[\mathcal{A}]$ of size m. Then,*

$$\lim_{m\to\infty} \frac{L[\mathcal{A}]^{[\geq k]}_{m,n}}{A_m} \sim \lim_{m\to\infty} \frac{L[\mathcal{A}]^{[k]}_{m,n}}{A_m} \sim \frac{d}{n^k},$$

when n tends to infinity, and d is a constant.

The proof of the generalisation works analogously to the one of Theorem 3.10, still we will state the main ideas as they will be useful in the following.

Proposition 3.13. *Let $\tilde{\mathcal{A}}$ be the family of associative trees with unlabelled leaves. Given an associative tree $t \in \tilde{L}[\tilde{\mathcal{A}}]_m$ with l L-pattern leaves. Further, we fix a set $\mathcal{V} \subseteq \mathcal{X}_n$ denote by v the cardinality of this set, $v = |\mathcal{V}|$. Then the number of leaf-labellings of t which make it have k $(L,\mathcal{V})$-restrictions is:*

$$(n-v)^{\underline{l-k}}\, n^{m-l}\, 2^m w_{v,k}(l),$$

where $w_{v,k}(l) = \sum_{r=0}^{k} \left\{ {l \atop l-r} \right\} \binom{v}{k-r} (l-r)^{\underline{k-r}}$.

Remark: Note that $w_{v,k}(l)$ is a polynomial in l.

Proof. For any $r \leq k$, the number of different leaf-labellings of t which give r L-repetitions and k $(L,\mathcal{V})$-restrictions is:

$$\left\{ {l \atop l-r} \right\} \binom{v}{k-r} (l-r)^{\underline{k-r}}\, (n-v)^{\underline{l-r-(k-r)}}\, n^{m-l} 2^m,$$

where $x^{\underline{y}} = x(x-1)\ldots(x-y+1)$ and $\left\{ {y \atop x} \right\}$ are the Stirling numbers of second kind In this formula, the different factors represent, from left to right:

- the number of partitions of the L-pattern leaves into $l-r$ classes (leaves in the same class will be labelled by the same variable),
- the number of different choices for the $k-r$ essential variables that appear in the L-pattern leaves,
- the number of different assignments of these essential variables to the $l-r$ classes of the first term,
- the number of assignments of non-essential variables to the remaining classes of the L-pattern leaves,
- the number of assignments of variables to the leaves that are not L-pattern leaves,
- the number of ways to distribute the negations. □

In [18] the following proposition is proved for binary trees and patterns (cf. [18, Lemma 2.7]), but in fact the proof does not rely on binarity and hence the proposition holds for patterns and trees of arbitrary degree.

Proposition 3.14. *Let $\mathcal{T}$ be a set of trees whose generating function $t(z) = \sum t_m z^m$ has a unique dominating singularity ρ in $\mathcal{R}^+$ of the square root type. Let $\tilde{L}$ be an unambiguous pattern language, subcritical for $\mathcal{T}$. Let $\tilde{L}[\mathcal{T}](m,l)$ denote the number of trees from $\tilde{L}[\mathcal{T}]$ of size m with exactly l pattern leaves. Finally, let $w(l)$ be a non zero polynomial of degree δ. Then,*

$$\lim_{n\to\infty} \frac{\sum_{l\geq 0} \tilde{L}[\mathcal{T}](m,l)w(l)}{t_m} = c_w,$$

for some non-negative real c_w.

Moreover, if $w(l)$ has non-negative values and is positive at some point l_0, and if L contains a pattern with l_0 non pattern leaves and at least one placeholder, then $c_w \neq 0$.

Thanks to those propositions, we can now prove Theorem 3.12 to associative trees:

Proof of Theorem 3.12. Let L be an associative pattern and $\tilde{A}$ the family of trees from $\mathcal{A}$ with unlabelled leaves. We have, thanks to Proposition 3.13:

$$\frac{L[\mathcal{A}]_{m,n}^{[k]}}{A_m} = \frac{2^m \sum_{l\geq 0} \tilde{L}[\tilde{\mathcal{A}}](m,l) w_{k,v}(l) (n-v)^{\underline{l-k}} n^{m-l}}{A_m};$$

and this implies:

$$\frac{L[\mathcal{A}]_{m,n}^{[k]}}{A_m} \leq \frac{2^m \sum_{l\geq 0} \tilde{L}[\tilde{\mathcal{A}}](m,l) w_{k,v}(l) n^{l-k} n^{m-l}}{(2n)^m \tilde{A}_m}.$$

Thanks to Proposition 3.14, we get

$$\lim_{m\to\infty} \frac{L[\mathcal{A}]_{m,n}^{[k]}}{A_m} \leq \lim_{m\to\infty} \frac{2^m \sum_{l\geq 0} L[\mathcal{A}](m,l) w_{k,v}(l) n^{m-k}}{(2n)^m \tilde{A}_m} \sim \frac{c_{k,v}}{n^k},$$

when n tends to infinity. Moreover, we can check that $c_{k,v}$ is positive. A lower bound can be proven analogously, the proof for the binary case is given in [18]. It follows that

$$\lim_{m\to\infty} \frac{L[\mathcal{A}]_{m,n}^{[k]}}{A_m} \sim \frac{d}{n^k},$$

when n tends to infinity. Moreover, we can see that

$$\frac{L[\mathcal{A}]_{m,n}^{[\geq k]}}{A_m} \leq \frac{2^m \sum_{l\geq 0} \tilde{L}[\tilde{\mathcal{A}}](m,l) w_{k,v}(l) n^{m-k}}{A_m},$$

and since

$$\lim_{m\to\infty} \frac{L[\mathcal{A}]_{m,n}^{[k]}}{A_m} \leq \lim_{m\to\infty} \frac{L[\mathcal{A}]_{m,n}^{[\geq k]}}{A_m},$$

the theorem is proven. □

3.2.2. Associative tautologies

Proposition 3.15. *In the associative model, asymptotically when n tends to infinity, almost all tautologies are simple tautologies.*

The proof relies on the ideas of the binary case (see [18]). First we need to introduce suitable pattern languages and show that they are subcritical for the tree family we are considering:

$$\begin{cases} \hat{N} = \bullet|\check{N} \wedge \Box|\check{N} \wedge \Box \wedge \Box| \dots \\ \check{N} = \bullet|\hat{N} \vee \hat{N}|\hat{N} \vee \hat{N} \vee \hat{N}| \dots \\ R = \hat{N}|\check{N} \end{cases} \tag{7}$$

Then R is an unambiguous pattern language.

Lemma 3.16. *The labelled pattern languages R and $R[R] := \hat{N}[\check{N}] \mid \check{N}[\hat{N}]$ are subcritical for associative trees.*[2]

Remark: Since the trees must be stratified, we cannot use the ordinary composition of pattern languages as defined in Definition 3.7. Therefore it is not *a priori* obvious that the modified composition of subcritical pattern languages is subcritical as well.

Proof. The generating function $p(x, y)$ of the labelled pattern language R is given by

$$p(x, y) = \hat{p}(x, y) + \check{p}(x, y) - 2nx,$$

where $\hat{p}(x, y)$ (resp. $\check{p}(x, y)$) is the generating function of the partial labelled patterns $\hat{N}$ (resp. $\check{N}$). These two generating functions satisfy the following system:

$$\begin{cases} \check{p}(x, y) = 2nx + \frac{\hat{p}(x,y)^2}{1-\hat{p}(x,y)} \\ \hat{p}(x, y) = 2nx + \frac{y}{1-y}\check{p}(x, y). \end{cases} \tag{8}$$

Solving this system, we get

$$\hat{p}(x, y) = \frac{1}{2}\left(2nx - y + 1 - \sqrt{(2nx - y + 1)^2 - 8nx}\right).$$

Recall (cf. (2)) that the generating function for associative trees is defined by

$$A(z) = \frac{1}{2}\left(1 - 2nz - \sqrt{1 - 12nz + 4n^2z^2}\right),$$

$$A(\alpha_n) = \sqrt{2} - 1 \text{ and } \alpha_n = \frac{3 - 2\sqrt{2}}{2n}$$

and for trees with root label *And* resp. *Or* by

$$\hat{A}(z) = \check{A}(z) = \frac{A(z) + 2nz}{2} \text{ and } \hat{A}(\alpha_n) = \check{A}(\alpha_n) = \frac{2 - \sqrt{2}}{2}.$$

By stratification, subcriticality means here that $\hat{p}(x, y)$ and $\check{p}(x, y)$ are analytic in some domain $\mathcal{D} = \{(x, y) \in \mathbb{C}^2 \mid |x| \leq \alpha_n + \epsilon, |y| \leq (2 - \sqrt{2})/2 + \epsilon\}$. Obviously, $\hat{p}(x, y)$ is analytic in $\mathcal{E} = \mathbb{C}^2 \setminus \{(x, y) \mid (2nx - 1 - y)^2 = 8nx\}$. Thus setting $x = \alpha_n z$ with $|z| \leq 1 + \epsilon$ and solving the equation defining $\mathcal{E}$ we obtain

$$y = 1 + (3 - 2\sqrt{2})z \pm 2(\sqrt{2} - 1)\sqrt{z} \approx 1 + 0.171573z \pm 0.828427\sqrt{z}.$$

[2] Strictly speaking, $R, \hat{N}, \check{N}$ are *unlabelled* pattern languages and should therefore be denoted by $\tilde{R}, \tilde{\hat{N}}, \tilde{\check{N}}$. But we avoid such notations and use $R, \hat{N}, \check{N}$ for both, the labelled as well as the unlabelled versions of the respective pattern languages, *cf.* remark after Definition 3.7.

The minimal modulus is attained if we take the minus sign and set $z = 1$: This gives $6 - 4\sqrt{2} \approx 0.343146$ which is larger than $\check{A}(\alpha_n) \approx 0.292893$. Thus $\mathcal{D} \subseteq \mathcal{E}$ and we are left with showing that $\check{p}(x, y)$ is analytic in $\mathcal{D}$ as well. By (8) the latter is true if $|\hat{p}(x, y)| < 1$ in $\mathcal{D}$. To see this, observe that $|\hat{p}(x, y)|$ attains its maximum if $x > 0, y > 0$. Since $\hat{p}(\alpha_n, \check{A}(\alpha_n)) = \check{A}(\alpha_n) < 1$ we obtain subcriticality of R for associative trees.

The pattern language $R[R]$ is subcritical if $\hat{p}(x, \check{p}(x, y))$ and $\check{p}(x, \hat{p}(x, y))$ are analytic in $\mathcal{D}$. In view of the considerations above this is an easy consequence of

$$\hat{p}(\alpha_n, \check{A}(\alpha_n)) = \check{A}(\alpha_n) \quad \text{and} \quad \check{p}(\alpha_n, \hat{A}(\alpha_n)) = \hat{A}(\alpha_n)$$

and $\check{A}(\alpha_n) = \hat{A}(\alpha_n) < 1$. □

Remark: The R-pattern has an interesting property: if we set all the R-pattern leaves of a tree to $False$, then the whole tree itself computes $False$. This can be checked by induction on the size of the tree. If the pattern is only a leaf, it returns $False$. If the root is an $\vee$-node, then all subtrees of the root are patterns returning $False$ by the induction hypothesis. If the root is an $\wedge$-node, the leftmost subtree is a pattern returning $False$ by the induction hypothesis. Thus the whole tree computes $False$ in all cases. This property is the key point of the following proof.

Remark: The pattern R is a generalisation of the pattern $N = \bullet|N \vee N|N \wedge \square$, defined in [18] to handle the proof in the binary plane case. Note that $R[\mathcal{A}] = \mathcal{A}$, and we can find the unique element from $L[\mathcal{A}]$ which corresponds to a tree $A \in \mathcal{A}$ by starting at the root of A and finding the pattern leaves by traversing the tree top-to-bottom.

Proof of Proposition 3.15. Let us consider a tautology t with exactly one $(R[R], \emptyset)$-restriction (cf. Definition 3.7). This restriction has to be a repetition, since a tautology does not contain essential variables. (Thus we have set $\mathcal{V} = \emptyset$ here.)

If the repetition is of the kind x/x, then we can assign all the R-pattern leaves to $False$, and with this assignment the whole tree computes $False$, which is impossible.

Thus the repetition has to be an $x/\bar{x}$ repetition. Let us first assume that the repetition does not appear among the R-pattern leaves. Thus we can assign all those leaves to $False$, and then the whole term computes $False$, which is impossible. Hence, the repetition must occur in the R-pattern leaves. Let us assume that there is a node ν labelled by $\wedge$ on one of the paths from the leaves labelled by x and $\bar{x}$ to the root of the tree. Then the subtree rooted at ν has shape $t_1 \wedge t_2 \wedge \ldots \wedge t_s$ with $s \geq 2$. Let us assume that x (or $\bar{x}$) appears in t_j. Then we can assign all the $R[R]$-pattern leaves of the other subtrees $(t_i)_{i \neq j}$, and all the $R[R]$-pattern leaves of the whole tree except those in the subtree rooted at ν to $False$. This makes the whole tree compute $False$, which is impossible.

Thus, x and $\bar{x}$ are linked to the root by an $\vee$-only-path. As the trees are stratified, the only possibility for t is to be a simple tautology. Thus every tree with exactly one $(R[R], \emptyset)$-restriction computing $True$ is a simple tautology.

Moreover, there are no trees computing $True$ without $(R[R], \emptyset)$-restrictions, and the number of trees computing $True$ with at least two $(R[R], \emptyset)$-restrictions is negligible in comparison to the number of simple tautologies by Theorem 3.12 which can be applied thanks to Lemma 3.16. □

We are now able to prove Theorem 3.11 by counting associative simple tautologies.

Proof of Theorem 3.11. Let $\widetilde{ST}^x(z)$ be the generating function counting the number of simple tautologies realised by x and such that x and $\bar{x}$ appear only once in the first generation (i.e. at depth 1). Then,

$$\widetilde{ST}^x(z) = z^2 \sum_{l \geq 2} l(l-1)(\hat{A}(z) - 2z)^{l-2}.$$

If x or $\bar{x}$ appear at least twice in the first generation, the tree has at least two $(R[R], \emptyset)$-restrictions, and the set of such trees is negligibly small compared to the set counted by $\widetilde{ST}^x$. Thus, asymptotically, when n tends to infinity, $\widetilde{ST}^x(z)$ counts the set of simple tautologies realised by x.

Finally, note that in view of Theorem 3.12, the assertion of Proposition 3.5 extends to the associative case. So a Maple computation giving

$$\lim_{z\to\alpha_n}\frac{G'(z)}{A'(z)}\sim\frac{51-36\sqrt{2}}{n}$$

completes the proof of Theorem 3.11. □

3.3. The binary commutative model

The generating function of binary commutative And/Or trees, $C(z)=\sum_m C_m z^m$, is given in (3). We denote by γ_n the dominant positive singularity of $C(z)$. To compute γ_n and $C(\gamma_n)$ we need to solve the system (see [6, Chapter 2] for details):

$$\begin{cases} y = 2nz + y^2 + C(z^2) \\ 1 = 2y. \end{cases} \tag{9}$$

$C(z^2)$ is analytic for $|z|\le\gamma_n$. We obtain $C(\gamma_n)=\frac{1}{2}$ and $\gamma_n=\frac{1}{8n}-\frac{C(z^2)}{2n}$. As $C(z)=2nz+\mathcal{O}(n^2z^2)$, by inserting into the equation we can further derive $\gamma_n=\frac{1}{8n}\left(1-\frac{1}{8n}\right)+\mathcal{O}\left(\frac{1}{n^3}\right)$. As we need more terms in some of our calculations, we do a more refined analysis with Maple and further obtain

$$\gamma_n=\frac{1}{8n}\left(1-\frac{1}{8n}+\frac{7}{256n^2}\right)+\mathcal{O}\left(\frac{1}{n^4}\right). \tag{10}$$

Theorem 3.17. *The limiting probability of the function True in the binary commutative model is given by*

$$\mathbb{P}^c_n(True)=\lim_{m\to\infty}\mathbb{P}^c_{m,n}(True)=\frac{385}{512n}+\mathcal{O}\left(\frac{1}{n^2}\right).$$

To prove the theorem we will extend the method of pattern languages of Kozik to the commutative case. We consider binary commutative trees, together with a *half-embedding*, that is certain branches of the tree will be plane and some will stay non-plane. We use the plane pattern language known from Section 3.1 given by

$$N=\bullet|N\vee N|N\wedge\square.$$

As N is plane, it is unambiguous for any tree family. A tree of $N[\mathcal{C}]$ is a "mobile", that is, the pattern-trees consisting of internal nodes and • and □-leaves are plane, while the trees substituted into the □-nodes are still non-plane trees.

3.3.1. Generalisation of Kozik's theorem to commutative trees

As mentioned before, in the plane case, the pattern N we considered fulfilled $N[\mathcal{T}]=\mathcal{T}$. For commutative trees, this is not the case. The proof of Theorem 3.10 relies completely on plane structures and subcriticality which is not given anymore. Still, Theorem 3.10 can be generalised to mobile structures. We will adapt it and its proof, relying on the sketch in Section 3.2.1, but we will need additional arguments. Note that most definitions concerning pattern languages, such as restrictions, can still be used without change. However, the pattern languages we encounter here are not subcritical any more. Therefore, we will need a different concept of subcriticality.

Definition 3.18. Let $f(x)$ and $g(x)$ be power series with nonnegative coefficients. Assume further that $g(x)$ has a unique singularity $\rho>0$ on its circle of convergence and that $g(\rho)=\eta<\infty$. We say that f is subcritical for g if f is analytic for $|x|\le\eta+\epsilon$ for some $\epsilon>0$.

Theorem 3.19. *Let L be a labelled plane binary unambiguous pattern language with $\ell(x,y)$ its generating function. Further assume that the coefficients $A_l(y)$, given by*

$$\ell(x,y)=\sum_{l\ge0}\sum_{i\ge0}s_{i,l}y^ix^l=\sum_{l\ge0}A_l(y)x^l \tag{11}$$

are subcritical for $C(z)$.

Fix a set $\mathcal{V} \subseteq \mathcal{X}_n$. We denote by $L[\mathcal{C}]^{[k]}_{m,n}$ (resp. by $L[\mathcal{C}]^{[\geq k]}_{m,n}$) the number of elements of $L[\mathcal{C}]$ of size m which have k (resp. at least k) $(L,\mathcal{V})$-restrictions, and by $L[\mathcal{C}]_m$ the number of elements of $L[\mathcal{C}]$ of size m. Then,

$$\lim_{m\to\infty} \frac{L[\mathcal{C}]^{[\geq k]}_m}{C_m} = \mathcal{O}\left(\frac{1}{n^k}\right) \text{ and } \lim_{m\to\infty} \frac{L[\mathcal{C}]^{[k]}_m}{C_m} = \mathcal{O}\left(\frac{1}{n^k}\right)$$

when n tends to infinity.

Let $\tilde{L}$ be a plane pattern and $\mathcal{C}$ be the family of commutative trees. Let Λ be an element of $\tilde{L}[\mathcal{C}]$ of size m with l pattern leaves. Note that the leaves of the non-plane parts are labelled while the pattern leaves are unlabelled.

Proposition 3.20. *Fix the set $\mathcal{V} \subseteq \mathcal{X}_n$ of essential variables and denote by $v = |\mathcal{V}|$ the cardinality of this set. Given a binary mobile $\Lambda \in L[\mathcal{C}]$ with unlabelled leaves, the number of leaf-labellings of Λ which make it have k $(L,\mathcal{V})$-restrictions satisfies*

$$\sharp(\textit{labellings})_k = (n-v)^{\underline{l-k}}\, 2^l w_{v,k}(l),$$

where the last factor is the polynomial $w_{v,k}(l) = \sum_{r=0}^{k} \left\{ {l \atop l-r} \right\} \binom{v}{k-r} (l-r)^{\underline{k-r}}$.

Proof. For any $r \leq k$, the number of different labellings of the pattern leaves of Λ which give r L-repetitions and k $(L,\mathcal{V})$-restrictions is given by

$$\left\{ {l \atop l-r} \right\} \binom{v}{k-r} (l-r)^{\underline{k-r}}\, (n-v)^{\underline{l-r-(k-r)}}\, 2^l,$$

where, as in the plane case, $x^{\underline{y}} = x(x-1)\dots(x-y+1)$ and $\left\{ {y \atop x} \right\}$ are the Stirling numbers of second kind. The different terms of the product again represent, from left to right:

- the number of partitions of the L-pattern leaves into $l-r$ classes (leaves in the same class will be labelled by the same variable),
- the number of different choices for the $k-r$ essential variables that appear in the L-pattern leaves,
- the number of different assignments of these essential variables to the $l-r$ classes of the first term,
- the number of assignments of non-essential variables to the remaining classes of the L-pattern leaves,
- and the number of distribution of the negations. □

We adapt Proposition 3.14 to our needs.

Proposition 3.21. *Let L be an unambiguous labelled pattern language, with $\ell(x,y)$ its generating function, and let $\mathcal{T}$ be a family of leaf-labelled trees with generating function $T(z)$. Further assume that the coefficients $A_l(y) := [x^l]\ell($, given in (11), are subcritical for $T(z)$.*

Let $L[\mathcal{T}](m,l)$ be the number of trees of $L[\mathcal{T}]$ of size m with exactly l pattern leaves and $w(l)$ be a non-zero polynomial of degree λ. Then,

$$\lim_{m\to\infty} \frac{\sum_{l=0}^{N} L[\mathcal{T}](m,l)w(l)}{T_m} = c_w, \tag{12}$$

for some non-negative real c_w, where N is some fixed integer.

Proof. If we set $\ell_w(x,y) = \sum_{l=0}^{N} w(l)A_l(y)x^l$, then the numerator of (12) can be expressed as $[z^n]\ell_w(z,C(z))$. Moreover, $w(l) = \sum_{j=0}^{\lambda} w_j l^{\underline{j}}$ is a representation of the polynomial w, and $\ell_N(x,y) = \sum_{l=0}^{N} A_l(y)x^l$ is the truncation of $\ell(x,y) = \sum_{l\geq 0} A_l(y)x^l$. Note that,

$$x^j \frac{\partial^j \ell_N(x,y)}{\partial x^j} = \sum_{l=0}^{N} l^{\underline{j}} A_l(y)x^l.$$

Thus

$$\sum_{j=0}^{\lambda} w_j x^j \frac{\partial^j \ell_N(x,y)}{\partial x^j} = \sum_{l=0}^{N} w(l)A_l(y)x^l.$$

Hence, the generating function $\ell_w(x,y)$ is a linear combination of $\ell_N(x,y)$ and its derivatives, all of which are finite sums of terms which are subcritical for $T(z)$. Hence, $\ell_w(z,C(z))$ and $T(z)$ have the same radius of convergence. By [18, Observation 2.3] every subcritical summand has a square root expansion around the singularity, if $T(z)$ has a square root singularity, hence the type of singularity of $\ell_w(z,C(z))$ is also of order $1/2$ or of higher order if there is a cancellation.

Thanks to a transfer lemma [8], we easily get

$$\frac{[z^m]\ell_w(z,C(z))}{[z^m]T(z)} \sim const,$$

when m tends to infinity. Therefore,

$$\lim_{m\to\infty} \frac{\sum_{l\geq 0} L[\mathcal{T}](m,l)w(l)}{T_m} = c_w$$

for some non-negative constant c_w. Further c_w is positive if there is no cancellation, and zero otherwise. ☐

Proof of Theorem 3.19. We have, thanks to Proposition 3.20:

$$\frac{L[\mathcal{C}]_m^{[k]}}{C_m} = \frac{\sum_{l=0}^{N} \tilde{L}[\mathcal{C}](m,l)w_{k,v}(l)(n-v)^{\underline{l-k}}\, 2^l}{C_m},$$

where $N = n - v + k$, because for larger l the factor $(n-v)^{\underline{l-k}}$ gives 0. This implies:

$$\frac{L[\mathcal{C}]_m^{[k]}}{C_m} \leq \frac{\sum_{l=0}^{N} \tilde{L}[\mathcal{C}](m,l)w_{k,v}(l)n^{l-k}\, 2^l}{C_m} = \frac{\sum_{l=0}^{N} L[\mathcal{C}](m,l)w_{k,v}(l)}{C_m} \cdot \frac{1}{n^k}.$$

because $L[\mathcal{C}](m,l) = (2n)^l \tilde{L}[\mathcal{C}](m,l)$. And therefore, by applying Proposition 3.21, we get that

$$\lim_{m\to\infty} \frac{L[\mathcal{C}]_m^{[k]}}{C_m} \leq \frac{c_w}{n^k}.$$

The proof for $\lim_{m\to\infty} \frac{L[\mathcal{C}]_m^{[\geq k]}}{C_m}$ is analogous to the latter one. ☐

3.3.2. Commutative tautologies

Proposition 3.22. *In the commutative model, asymptotically almost all tautologies are simple tautologies when n tends to infinity.*

Before proving Proposition 3.22, we introduce some *half-embedding* of a tree t into the plane: Start at the root and choose a left-right order of the children of the root. If the root is an $\wedge$-node, proceed recursively with the root of the left subtree, the right subtree remains non-plane. If the root is an $\vee$-node, proceed recursively with both subtrees. If doing so we meet a leaf, it is a pattern leaf. Doing this for the whole tree t, we obtain an element of $N[\mathcal{C}]$, where the non-plane subtrees are the structures substituted into the placeholders. Now do the same half-embedding starting at every root of a non-plane subtree. Thus we obtain an element of $N[N][\mathcal{C}]$. Note that different trees $t_1 \neq t_2 \in \mathcal{C}$ will create different patterns $N[t_1]$ and $N[t_2]$, thus the function $\mathcal{C} \to N[\mathcal{C}]$ described above is an injection. Of course, there are several ways to embed a tree t with the above method, so for every tree t we choose an embedding such that the resulting $N[N]$-pattern has a minimal number of $(N[N], \mathcal{V})$-restrictions. We call such an embedding a minimal $N[N]$-embedding of t (of course there could be various minimal embeddings for one tree).

Lemma 3.23. *Let t be a tree computing the function $True$. Then its minimal $[N]$-embeddings have at least one $(N, \emptyset)$-restriction.*

Proof. Suppose $N[t]$ has no restriction and set all pattern leaves to $False$. We proceed inductively. If $N[t]$ is just a leaf, it returns $False$. If the root of $N[t]$ is an $\wedge$-node, the left subtree is a pattern and will, by the induction hypothesis, return $False$, thus the whole tree returns $False$. If the root of $N[t]$ is an $\vee$ node, both subtrees are patterns returning $False$ by the induction hypothesis. Thus the whole tree returns $False$. □

Lemma 3.24. *Let t be a tree whose minimal $N[N]$-embeddings have exactly one $(N[N], \emptyset)$-restriction. Then t is a simple tautology.*

Proof. There are two cases to distinguish.

First case:. The restriction is of type x/x. Set all N-pattern leaves to $False$. The same arguments as in the proof of Lemma 3.23 show that t returns $False$.

Second case:. The restriction is of type $x/\bar{x}$. Then the restriction appears on the first level, that is, in $N[t]$, as otherwise setting all N-pattern leaves to $False$ would lead to a tree computing $False$ by the same arguments as before. If t is not a simple tautology, then there exists at least one node labelled with $\wedge$ on the path from the root to either x or its negation. Let t_1 be the non-plane subtree rooted at such a node. After the second N-embedding, the $N[t_1]$ pattern contains no repetition as the whole tree $N[N][t]$ had only one $N[N]$-repetition, thus it is easy to have t_1 contribute $False$ by setting all $N[t_1]$-pattern leaves to $False$. Then the $\wedge$-node at ν gives $False$, thus t does not compute the function $True$. Hence, every tautology t which has a minimal $N[N]$-embedding with a single repetition is a simple tautology. □

Lemma 3.25. *Let L be a pattern language with generating function $\ell(x,y) = \sum_{l \geq 0} A_l(y) x^l$ and with $A_0(y) = 0$, and let L^r be its r-th power for any $r \in \mathbb{N}$, with*

$$\ell^*(x,y) = \underbrace{\ell(x, (\ell(x, \cdots \ell(x}_{r \text{ times}}, y) \cdots))) = \sum_{l \geq 0} A_l^*(y) x^l$$

its generating function. Further let $\mathcal{T}$ be a family of trees with generating function $T(z)$. Assume that, for all $l \geq 0$, $A_l(y)$ is subcritical for $T(z)$. Then $A_l^*(y)$ is subcritical for $T(z)$.

Proof. First note that $A_0(y) = 0$ means that every pattern in L has at least one pattern leaf. Obviously, this property still holds for $A_0^*(y)$.

We prove the statement by induction: the case $r = 1$ is true by assumption. Let us assume that the result holds for r, and let $\bar{s}(x,y) = \sum_{l \geq 0} \bar{A}_l(y) x^l$ be the generating function of L^r with $\bar{A}_l(y)$ being subcritical for $T(z)$. We want to show that $[x^l] s(x, \bar{s}(x,y))$ is subcritical for $T(z)$. It is sufficient to show that $[x^\lambda] A_l(\bar{s}(x,y))$ is subcritical for $T(z)$ for all λ, because $s(x, \bar{s}(x,y)) = \sum_{l \geq 0} A_l(\bar{s}(x,y)) x^l$ and $A_l(\bar{s}(x,y))$ is a power series in x, i.e. $[x^l] s(x, \bar{s}(x,y)) =$

$\sum_{j=0}^{l}[x^{l-j}]A_j(\bar{s}(x,y))$, which is a finite sum of such coefficients.

$$\begin{aligned}
[x^\lambda]A_l(\bar{s}(x,y)) &= [x^\lambda]\sum_{j\geq 0} s_{l,j}\bar{s}(x,y)^j \\
&= [x^\lambda]\sum_{j\geq 0} s_{l,j}\left(\sum_{\mu\geq 0} x^\mu \bar{A}_\mu(y)\right)^j \\
&= [x^\lambda]\sum_{j\geq 0} s_{l,j}\sum_{\mu_1,\ldots,\mu_j} x^{\sum \mu_i}\bar{A}_{\mu_1}(y)\cdots\bar{A}_{\mu_j}(y) \\
&= \sum_{j\geq 0} s_{l,j}\sum_{\mu_1+\ldots+\mu_j=\lambda}\bar{A}_{\mu_1}(y)\cdots\bar{A}_{\mu_j}(y).
\end{aligned}$$

As $\bar{A}_0(y)=0$, $\mu_i>0$ for $i=1,\ldots j$, and hence we have a maximum of λ factors in every summand, that is,

$$[x^\lambda]A_l(\bar{s}(x,y)) = \sum_{j=0}^{\lambda} s_{l,j}\sum_{\substack{\mu_1,\ldots,\mu_i,\\ \mu_1+\ldots+\mu_j=\lambda}}\bar{A}_{\mu_1}(y)\cdots\bar{A}_{\mu_j}(y).$$

This is a finite sum of finite products of subcritical factors and hence it is subcritical for $T(z)$. □

Lemma 3.26. *Let $s(x,y)=\sum_{l\geq 0}A_l(y)x^l$ be the generating function of the pattern N. The functions $A_l(y)$ are subcritical for $C(z)$.*

Proof. Thanks to symbolic arguments and the recursive definition of $N=\bullet|N\vee N|N\wedge\Box$ we get

$$s(x,y)=2nx+s(x,y)^2+ys(x,y).$$

Solving this equation gives $s(x,y)=\frac{1}{2}\left(1-y-\sqrt{(y-1)^2-8nx}\right)$. Thus we can deduce an explicit formula for the $A_l(y)$ from this expression.

$$\begin{aligned}
s(x,y) &= \frac{1-y}{2}-\frac{1}{2}\sqrt{(y-1)^2}\sqrt{1-\frac{8nx}{(y-1)^2}} \\
&= \frac{1-y}{2}-\frac{1}{2}(1-y)\sum_{l\geq 0}\binom{1/2}{l}(-8n)^l(y-1)^{-2l}x^l,
\end{aligned}$$

since $s(0,0)=0$. Therefore, $A_l(y)=-\frac{1}{2}(1-y)\binom{1/2}{l}(-8n)^l(y-1)^{-2l}$ is a rational function in y and its radius of convergence is 1. Hence, these functions are subcritical for $C(z)$. □

Proof of Proposition 3.22. Let t be a tree in $\mathcal{C}$ which computes *True*. Then there is at least one variable x appearing twice in the leaves of t, because otherwise the tree cannot be a tautology (induction on the size of the tree). We half-embed t into the plane as described before. As this N-embedding represents an injection it follows that $C_m^{[k]}\leq(N[\mathcal{C}])_m^{[k]}$, where $C_m^{[k]}$ denotes the number of trees from $\mathcal{C}$ of size m whose minimal half-embeddings have k restrictions. Hence, by Theorem 3.19, which can be applied thanks to Lemmas 3.25 and 3.26:

$$\frac{C_m^{[k])}}{C_m}\leq\frac{N[\mathcal{C}]_m^{[k])}}{C_m}=\mathcal{O}\left(\frac{1}{n^k}\right),$$

and thus asymptotically almost all tautologies in a binary commutative And/Or tree are simple (and have a minimal $N[N]$-embedding with one restriction). Proposition 3.22 is thus proved. □

Proof of Theorem 3.17. Let $g_x(z)$ be the generating function counting the trees in $\mathcal{C}$ with a $\vee$-only-path from the root to a leaf labelled with x. It is given by $g_x(z) = C(z) - \bar{g}_x(z)$ with

$$\bar{g}_x(z) = (2n-1)z + \frac{1}{2}\left(C^2(z) + C(z^2)\right) + \frac{1}{2}\left(\bar{g}_x^2(z) + \bar{g}_x(z^2)\right), \tag{13}$$

because a tree rooted at an $\wedge$-node cannot contain an $\vee$-only-path from the root, while if the root is labelled with $\vee$ both subtrees of the root must not contain an $\vee$-only-path to an x-leaf.

The generating function $ST^x(z)$ of the family of trees which are a simple tautology realized by x is given by $ST^x(z) = C(z) - \overline{ST}^x(z)$, where $\overline{ST}^x(z)$ corresponds to the family of trees which are not simple tautologies realised by x. Similarly to $\bar{g}_x(z)$, such a tree is either rooted at an $\wedge$-node, or it is rooted at an $\vee$-node, and both subtrees of the root are not simple tautologies. Still, it could return $True$ if one of the subtrees contains an $\vee$-only-path to x and the other subtree contains an $\vee$-only-path to $\bar{x}$. The case where the subtrees are of this shape and one of them is even a tautology is already excluded by construction and therefore must not be subtracted. This gives the following implicit equation for $\overline{ST}^x(z)$.

$$\begin{aligned}\overline{ST}^x(z) &= 2nz + \frac{C^2(z)+C(z^2)}{2} + \frac{(\overline{ST}^x)^2(z) + \overline{ST}^x(z^2)}{2} - (g_x(z) - ST^x(z))(g_{\bar{x}}(z) - ST^x(z)),\\ &= 2nz + \frac{C^2(z)+C(z^2)}{2} + \frac{(\overline{ST}^x)^2(z) + \overline{ST}^x(z^2)}{2} - (\overline{ST}^x(z) - \bar{g}_x(z))^2. \end{aligned}\tag{14}$$

To calculate the limiting ratio of simple tautologies, we need to determine $n(1 - \lim_{z\to\gamma_n} \frac{(\overline{ST}^x)'(z)}{C'(z)})$, where the factor n is the choice of x in the set of variables, and we use an analogue of Proposition 3.5 as well as Lemma 3.4. We denote by $u_n := \bar{g}_x(\gamma_n)$, $v_n := \bar{g}(\gamma_n^2), U_n := \overline{ST}^x(\gamma_n)$ and $V_n := \overline{ST}^x(\gamma_n^2)$, and compute U_n up to terms of order $\frac{1}{n^2}$. From (13) we get

$$u_n = (2n-1)\frac{1}{8n}\left(1 - \frac{1}{8n}\right) + \frac{1}{2}\left(\frac{1}{4} + C(\gamma_n^2)\right) + \frac{1}{2}(u_n^2 + v_n) + \mathcal{O}\left(\frac{1}{n^2}\right) \tag{15}$$

$$v_n = (2n-1)\frac{1}{64n^2}\left(1 - \frac{1}{8n}\right)^2 + \frac{1}{2}(C^2(\gamma_n^2) + C(\gamma_n^4)) + \frac{1}{2}(v_n^2 + \bar{g}_x(\gamma_n^4)) + \mathcal{O}\left(\frac{1}{n^2}\right) \tag{16}$$

We know that $C(z^2) = 2nz^2 + \mathcal{O}(n^2z^4)$, hence $C(\gamma_n^2) = \frac{1}{32n} - \frac{1}{128n^2} + \mathcal{O}(\frac{1}{n^3})$. Inserting this into (16) we can compute $v_n = \frac{1}{32n} + \mathcal{O}\left(\frac{1}{n^2}\right)$, and with (15), we compute $u_n = \frac{1}{2} - \frac{1}{4n} + \mathcal{O}\left(\frac{1}{n^2}\right)$. Solving the equations for U_n and V_n, up to terms of order $\frac{1}{n^2}$, we get

$$V_n = \frac{1}{32n} - \frac{7}{1024n^2} + \mathcal{O}\left(\frac{1}{n^3}\right) \quad \text{and} \quad U_n = \frac{1}{2} - \frac{129}{1024n^2} + \mathcal{O}\left(\frac{1}{n^3}\right).$$

Differentiating $\overline{ST}^x(z)$ and $\bar{g}_x(z)$, we obtain

$$\begin{aligned}\bar{g}_x'(z) &= 2n - 1 + C(z)C'(z) + zC'(z^2) + \bar{g}_x(z)\bar{g}_x'(z) + z\bar{g}_x'(z^2),\\ (\overline{ST}^x)'(z) &= 2n + C(z)C'(z) + zC'(z^2) + \overline{ST}^x(z)\bar{g}_x'(z)\\ &\quad + z(\overline{ST}^x)'(z^2) - 2(\overline{ST}^x(z) - \bar{g}_x(z))((\overline{ST}^x)'(z) - \bar{g}_x'(z))\end{aligned}$$

Hence, recalling (9) and $C(\gamma_n) = 1/2$, we obtain

$$\begin{aligned}\lim_{z\to\gamma_n}\frac{\bar{g}_x'(z)}{C'(z)} &= \lim_{z\to\gamma_n}\frac{1}{1-\bar{g}_x(z)}\left(\frac{2n-1}{C'(z)} + C(z) + \frac{zC'(z^2)}{C'(z)} + \frac{z\bar{g}_x'(z^2)}{C'(z)}\right)\\ &\sim \frac{1}{2(1-u_n)} = 1 - \frac{1}{2n} + \frac{1}{4n^2} + \mathcal{O}(\frac{1}{n^3}),\end{aligned}$$

and

$$
\begin{aligned}
X_n &:= \lim_{z\to\gamma_n} \frac{(\overline{ST}^x)'(z)}{C'(z)} \\
&= \lim_{z\to\gamma_n} \frac{1}{1-\overline{ST}^x(z)} \\
&\quad\times\left(\frac{2n}{C'(z)} + C(z) + \frac{zC'(z^2)}{C'(z)} + \frac{z(\overline{ST}^x)'(z^2)}{C'(z)} - \frac{2(\overline{ST}^x(z)-\bar g_x(z))((\overline{ST}^x)'(z)-\bar g'_x(z))}{C'(z)}\right) \\
&\sim \frac{1}{1-U_n}\left(\frac{1}{2} - 2(U_n-u_n)\lim_{z\to\gamma_n}\frac{(\overline{ST}^x)'(z)-\bar g'_x(z)}{C'(z)}\right) \\
&\sim \left(2-\frac{129}{256n^2}\right)\left(\frac{1}{2}+\frac{1-X_n}{2n}\right).
\end{aligned}
$$

Solving the last asymptotic equivalence gives $X_n = 1 - \frac{385}{512n^2} + \mathcal{O}\left(\frac{1}{n^3}\right)$ and the result of Theorem 3.17 follows immediately. ☐

3.4. The associative and commutative model

The generating function of associative commutative And/Or trees $P(z)$ is given in (4) and (5). Note that $\hat{P}(z) = \check{P}(z)$. Let δ_n be the dominant positive singularity of $\hat{P}(z)$, and hence also of $P(z)$. To get $\delta_n, \hat{P}(\delta_n)$ and $P(\delta_n)$ we need to solve the system

$$
\begin{cases} y = e^y \cdot \Pi(z) - 1 - y + 2nz \\ 1 = e^y \cdot \Pi(z) - 1, \end{cases}
$$

with $\Pi(z) = \exp(\sum_{i\geq 2}\hat{P}(z^i)/i) = 1 + nz^2 + \mathcal{O}(nz^3)$, (since $\hat{P}(z) = 2nz + \mathcal{O}(n^2z^2)$). Therefore $\Pi(z) \sim 1$ for $z = \mathcal{O}\left(\frac{1}{n}\right)$ and n tending to infinity, hence the second equation gives $e^{y(z)} \sim 2$ or $y(z) \sim \ln(2)$. Inserting this value into the first equation gives $y = \frac{1+2nz}{2}$ and thus the first order asymptotic of δ_n is $\delta_n \sim \frac{2\ln 2 - 1}{2n}$.

Theorem 3.27. *The limiting probability of the function True in the associative and commutative model is given by*

$$
\mathbb{P}_n^{a,c}(True) = \lim_{m\to\infty}\mathbb{P}_{n,m}^{a,c}(True) = \frac{(2\ln 2-1)^2}{4n} + \mathcal{O}\left(\frac{1}{n^2}\right).
$$

To prove the theorem we will again use mobiles, using the unambiguous pattern $R = \hat{N}|\check{N}$ from Section 3.2, given in (7). We can prove that its coefficients $A_l(y)$ are subcritical for $\mathcal{P}$.

Lemma 3.28. *Let $p(x,y) = \sum_{l\geq 0} A_l(y)x^l$ being the generating function of the pattern language R. The functions $A_l(y)$ are subcritical for $P(z)$.*

Proof. The generating function of the R pattern is $p(x,y) = \hat{p}(x,y) + \check{p}(x,y) - 2nx$ where

$$
\hat{p}(x,y) = \frac{1}{2}\left(2nx - y + 1 - \sqrt{(2nx-y+1)^2 - 8nx}\right), \quad \text{and} \quad \check{p}(x,y) = 2nx + \frac{\hat{p}(x,y)^2}{1-\hat{p}(x,y)}.
$$

Thus,

$$
\begin{aligned}
\hat{p}(x,y) &= \frac{2nx-y+1}{2} - \frac{1-y}{2}\sum_{l\geq 0}\binom{1/2}{l}\frac{(4nx)^l(nx-1-y)^l}{(1-y)^{2l}} \\
&= \frac{2nx-y+1}{2} + \sum_{l\geq 0} A_\ell(y)x^l
\end{aligned}
$$

where the $A_l(y)$ are rational functions which are analytic at $y = 0$ and have radius of convergence equal to 1. Since $1 > \hat{P}(\delta_n) \sim \ln(2)$, $A_l(y)$ is subcritical for $P(z)$. ☐

3.4.1. Generalisation of Kozik's theorem to associative and commutative trees

Theorem 3.29. *Fix a set $\mathcal{V} \subseteq \mathcal{X}_n$ and let L be a labelled unambiguous pattern language where all nodes have out-degree different from 1. Further assume that the coefficients $A_l(y)$, given in* (11), *are subcritical for $P(z)$.*

We denote by $L[\mathcal{P}]^{[k]}_{m,n}$ (resp. by $L[\mathcal{P}]^{[\geq k]}_{m,n}$) the number of elements of $L[\mathcal{P}]$ of size m which have k (resp. at least k) $(L, \mathcal{V})$-restrictions. Then,

$$\lim_{m\to\infty} \frac{L[\mathcal{P}]^{[\geq k]}_{m,n}}{P_m} = \mathcal{O}\left(\frac{1}{n^k}\right) \text{ and } \lim_{m\to\infty} \frac{L[\mathcal{P}]^{[k]}_{m,n}}{P_m} = \mathcal{O}\left(\frac{1}{n^k}\right),$$

when n tends to infinity.

The proof of Theorem 3.29 now is an easy generalisation of Sections 3.2.1 and 3.3.1. We use mobiles on a associative plane pattern L, that is pattern leaves are on plane paths from the root, while commutative trees have been substituted in the □-nodes of the plane pattern. We can easily extend Proposition 3.20.

Proof of Theorem 3.29. As in previous sections, Proposition 3.20 gives:

$$\frac{L[\mathcal{P}]^{[k]}_{m,n}}{P_m} = \frac{\sum_{l\in\mathbb{N}} \tilde{L}[\mathcal{P}](m,l) w_{k,v}(l)(n-v)^{\underline{l-k}} 2^l}{P_m};$$

which implies:

$$\frac{L[\mathcal{P}]^{[k]}_{m,n}}{P_m} \leq \frac{\sum_{l\in\mathbb{N}} \tilde{L}[\mathcal{P}](m,l) w_{k,v}(l) n^{l-k} 2^l}{P_m} = \frac{\sum_{l\in\mathbb{N}} L[\mathcal{P}](m,l) w_{k,v}(l)}{P_m}.$$

Hence the result follows from Proposition 3.21. □

3.4.2. Non-plane associative tautologies

Proposition 3.30. *In the associative and commutative model, asymptotically almost all tautologies are simple tautologies, when n tends to infinity.*

Again we introduce a *half-embedding* of a tree t into the plane: Start at the root and choose a left to right order of the children of the root. If the root is an $\wedge$-node, proceed with the leftmost child of the root. If the root was an $\vee$-node, then do the same for every child of the root. If we end up at a leaf, this is a pattern leaf. By this procedure we obtain an element of $R[\mathcal{P}]$. Applying the same procedure to every root of a commutative subtree, we obtain an element of $R[R][\mathcal{P}]$, we call it an $R[R]$-embedding of t. There are several ways to embed t, choose one embedding with a minimal number of $(R[R], \emptyset)$-restrictions. Again, the function $t \mapsto R[R]_{min}(t)$ represents an injection.

Now looking at all trees with a minimal $R[R]$-embedding having exactly one restriction, we can proceed in the same way as in the proof of Theorem 3.11 to prove that they are simple tautologies.

Proof of Proposition 3.30. Let $t \in \mathcal{P}$ be a tree that computes True. We half-embed t and argue as in Section 3.3.2 to prove Proposition 3.30 with the help of Theorem 3.29. □

Proof of Theorem 3.27. We define $ST^x(z)$ as previously and obtain

$$\begin{aligned} ST^x(z) &= z^2 \sum_{\ell\geq 0} Z_\ell((\hat{P}(z) - 2z, \hat{P}(z^2) - 2z^2, \ldots) \\ &= z^2 \exp\left(\sum_{\ell\geq 1} \frac{\hat{P}(z^\ell) - 2z^\ell}{\ell}\right), \end{aligned} \tag{17}$$

where $Z_\ell(s_1, s_2, \dots)$ denotes the cycle index of the symmetric group on ℓ elements (c.f. [22]). Hence

$$\begin{aligned} ST^x(z) &= 2z\exp\left(\sum_{\ell\geq 1}\frac{\hat{P}(z^\ell)-2z^\ell}{\ell}\right) + z^2\exp\left(\sum_{\ell\geq 1}\frac{\hat{P}(z^\ell)-2z^\ell}{\ell}\right)\left(\sum_{\ell\geq 1} z^{\ell-1}(\hat{P}'(z^i)-2)\right) \\ &= \frac{2}{z}ST^x(z) + ST^x(z)\left(\hat{P}'(z) - 2 + \sum_{\ell\geq 2} z^{\ell-1}(\hat{P}'(z^i)-2)\right) \end{aligned}$$

At $z=\delta_n \sim \frac{2\ln 2-1}{2n}$, $ST^x(z)$ equals

$$ST^x(\delta_n) = \delta_n^2 \underbrace{\exp\left(\sum_{i\geq 1}\frac{\hat{P}(\delta_n^i)}{i}\right)}_{=2}\underbrace{\exp\left(\sum_{i\geq 1}\frac{-2\delta_n^i}{i}\right)}_{=(1-\delta_n)^2\sim 1} \sim 2\delta_n^2,$$

Hence, due to $P(z) = 2\hat{P}(z) - 2nz$,

$$\begin{aligned} \lim_{z\to\delta_n}\frac{G'(z)}{P'(z)} &= n\lim_{z\to\delta_n}\frac{ST^x(z)\hat{P}'(z)}{P'(z)} \\ &= n\lim_{z\to\delta_n}\frac{ST^x(z)\hat{P}'(z)}{2\hat{P}'(z)-2n} \sim \frac{(2\ln 2-1)^2}{4n}. \end{aligned}$$

□

4. Limiting probability of literals

In this section, we will compute the limiting probabilities of functions of complexity $L(f)=1$, that are literals x or $\bar{x}$. Therefore, in analogy to Section 3, we will define so called simple x-trees.

Definition 4.1. A simple x is a tree of the shape $x \wedge ST$, $x \vee SC$, $x \wedge (x \vee \cdots)$ or $x \vee (x \wedge \cdots)$, where ST denotes a simple tautology and SC a simple contradiction. The shape of such trees is depicted in Figures 6 and 7.

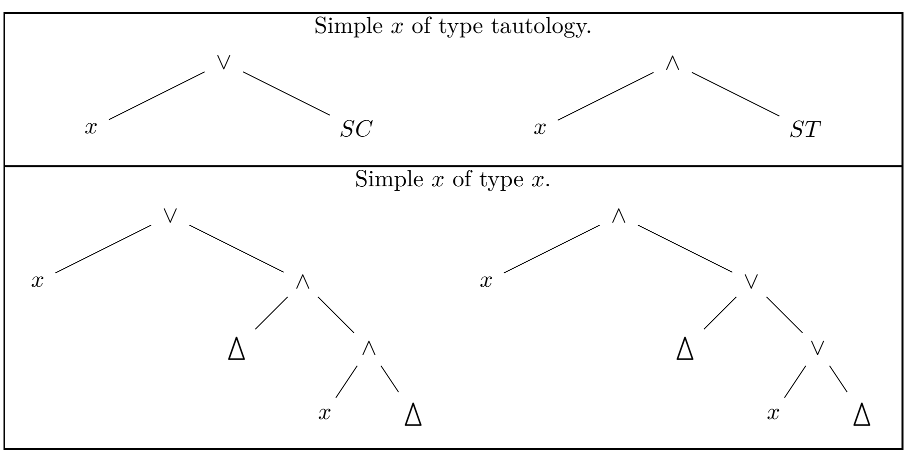


Figure 6: The different kinds of simple x. Here, ST is a simple tautology and SC is a simple contradiction.

For all models, we will prove the following proposition:

Proposition 4.2. *Asymptotically, almost all trees computing the function x are simple x.*

We state the proposition without a complete proof. The proof is easily done by similar arguments as in the previous section, using the patterns $N[N]$ or $R[R]$, respectively. We can prove that every tree $t \in \mathcal{T}$ or $t \in \mathcal{C}$ with exactly two $(N[N], \{x\})$-restrictions computing x, and every tree $t \in \mathcal{A}$ or $t \in \mathcal{P}$ with exactly two $(R[R], \{x\})$-restrictions, respectively, computing x is a simple x tree. Theorem 3.10 and its generalizations imply that those trees give asymptotically almost all trees computing x, as it is an easy task to prove that a large tree computing x will have at least two restrictions. Still we suggest a much simpler argument which proves the proposition in Section 6.

4.1. Binary plane trees.

Theorem 4.3. *The limiting probability of functions of complexity* 1 *in the binary plane model is*

$$\lim_{m\to\infty} \mathbb{P}_{m,n}(x) = \frac{5}{16n^2} + \mathcal{O}\left(\frac{1}{n^3}\right).$$

Proof. We distinguish between simple x of type tautology, which we denote by x_T, and simple x of type x, denoted by x_X (c.f. Figure 6). By Proposition 4.2, we have $\mathbb{P}_n(x) = \mathbb{P}_n(x_T) + \mathbb{P}_n(x_X)$.

First we compute $\mathbb{P}_n(x_T) = \lim_{m\to\infty} \mathbb{P}_{m,n}(x_T)$. Let $ST(z)$ be the generating function computing simple tautologies, given in Section 3.1. Of course, $ST(z)$ also counts contradictions. The generating function $\widetilde{ST}(z)$ of simple x of the first kind is given by $4z \cdot ST(z)$, where the factor z counts the leaf labelled with x, and the factor 4 is explained by the constant being a tautology or a contradiction, the label of the internal node then being fixed, and the constant being positioned left or right. Hence

$$\frac{[z^m]\widetilde{ST}(z)}{[z^m]T(z)} \sim 4\rho_n \frac{[z^m]ST(z)}{[z^m]T(z)} = 4\rho_n \mathbb{P}_n(True) \sim \frac{3}{16n^2}.$$

For the computation of $\mathbb{P}_n(x_X)$ we use the function $g_x(z)$ given in (6). Let $\tilde{g}_x(z)$ be the function counting simple x of type x. Then $\tilde{g}_x(z) = 4zg_x(z)$ by the same arguments as above, hence

$$\frac{[z^m]\tilde{g}_x(z)}{[z^m]T(z)} \sim 4\rho_n \frac{[z^m]g_x(z)}{[z^m]T(z)} \sim \frac{4}{16n} \lim_{z\to\frac{1}{16n}} \frac{g'(z)}{T'(z)}.$$

Using Maple, we get $\lim_{z\to\frac{1}{16n}} \frac{g'(z)}{T'(z)} = \frac{1}{2n} + \mathcal{O}\left(\frac{1}{n^2}\right)$, hence

$$\mathbb{P}_n(x) = \mathbb{P}_n(x_T) + \mathbb{P}_n(x_X) = \frac{3}{16n^2} + \frac{1}{8n^2} + \mathcal{O}\left(\frac{1}{n^3}\right) = \frac{5}{16n^2} + \mathcal{O}\left(\frac{1}{n^3}\right). \qquad \square$$

4.2. Associative plane trees.

Theorem 4.4. *The limiting probability of functions of complexity* 1 *in the associative model is*

$$\lim_{m\to\infty} \mathbb{P}^a_{m,n}(x) = \frac{546 - 386\sqrt{2}}{n^2} + \mathcal{O}\left(\frac{1}{n^3}\right).$$

Proof. Again we distinguish between simple x of type tautology (x_T), and simple x of type x (x_X, cf Figure 7). Note that a simple x in the associative case is represented by a tree with a binary root.

Calculating $\mathbb{P}_n(x_T) = \lim_{m\to\infty} \mathbb{P}_{m,n}(x_T)$, we obtain $\widetilde{ST}(z) = 4z\cdot ST(z)$ by the same arguments as above and

$$\frac{[z^m]\widetilde{ST}(z)}{[z^m]A(z)} \sim 4\alpha_n \frac{[z^m]ST(z)}{[z^m]A(z)} = 4\alpha_n \mathbb{P}^a_n(True) \sim 4\,\frac{3-2\sqrt{2}}{2n}\,\frac{51-36\sqrt{2}}{n} = \frac{594-420\sqrt{2}}{n^2}.$$

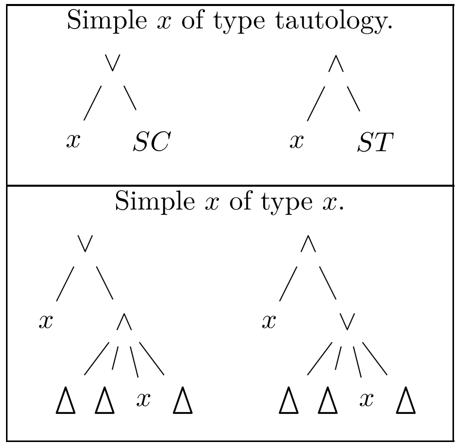


Figure 7: The different kinds of simple x in the associative case, up to commutativity. ST is a simple tautology and SC a simple contradiction.

The contribution of x_X is counted by $\tilde{g}(z) = 4zg(z)$, where $g(z)$ counts trees with an $\vee$-root and exactly one leaf labelled by x. Note that the other leaves may not be labelled with x neither with $\bar{x}$, because this would give a simple tautology. Then $g(z)$ is given by

$$g(z) = z \sum_{\ell \geq 2} \ell (A(z) - 2z)^{\ell - 1}.$$

Maple computations give $\lim_{z \to \frac{3-2\sqrt{2}}{2n}} \frac{g'(z)}{A'(z)} \sim \frac{3\sqrt{2}-4}{n} + \mathcal{O}\left(\frac{1}{n^2}\right)$, and thus

$$\frac{[z^m]\tilde{g}_x(z)}{[z^m]A(z)} \sim 4\alpha_n \frac{[z^m]g_x(z)}{[z^m]A(z)} \sim 4\,\frac{3-2\sqrt{2}}{2n}\,\frac{3\sqrt{2}-2}{n} = \frac{34\sqrt{2}-48}{n^2} + \mathcal{O}\left(\frac{1}{n^3}\right).$$

Adding the two limiting ratios gives the constant in Theorem 4.4. □

4.3. Binary commutative trees.

Theorem 4.5. *The limiting ratio of functions of complexity* 1 *in the binary commutative model is*

$$\lim_{m\to\infty} \mathbb{P}^c_{m,n}(x) = \frac{641}{2048n^2} + \mathcal{O}\left(\frac{1}{n^3}\right).$$

Proof. Simple x-trees are the same as in the plane binary case, but there is no left-to-right order anymore. Hence, $\widetilde{ST}(z) = 2\gamma_n ST(z)$, and $ST(z) = C(z) - \overline{ST}(z)$ with $\overline{ST}(z)$ given in (14). Hence

$$\frac{[z^m]\widetilde{ST}(z)}{[z^m]C(z)} \sim 2\gamma_n \frac{[z^m]ST(z)}{[z^m]C(z)} = 2\gamma_n \mathbb{P}^c_n(True) \sim 2\,\frac{1}{8n}\left(1 - \frac{1}{8n}\right)\frac{385}{512n} = \frac{385}{2048n^2} + \mathcal{O}\left(\frac{1}{n^3}\right),$$

$\tilde{g}_x(z) = 2\gamma_n g_x(z)$, and $g(z) = C(z) - \bar{g}_x(z)$ with $\bar{g}(z)$ given in (14) and $\lim \frac{\bar{g}'_x(z)}{C'(z)}$ computed in the proof of Theorem 3.17. Hence

$$\frac{[z^m]\tilde{g}_x(z)}{[z^m]C(z)} \sim 2\gamma_n \frac{[z^m]g_x(z)}{[z^m]C(z)} \sim 2\,\frac{1}{8n}\left(1 - \frac{1}{8n}\right)\frac{1}{2n} = \frac{1}{8n^2} + \mathcal{O}\left(\frac{1}{n^3}\right) = \frac{256}{2048n^2} + \mathcal{O}\left(\frac{1}{n^3}\right).$$ □

4.4. Associative and commutative trees.

Theorem 4.6. *The limiting ratio of functions of complexity* 1 *in the associative and commutative model is*

$$\lim_{m\to\infty} \mathbb{P}^{a,c}_{m,n}(x) = \frac{(2\ln 2 - 1)^2 (2\ln 2 + 1)}{4n^2} + \mathcal{O}\left(\frac{1}{n^3}\right).$$

Proof. Again, $\widetilde{ST}(z) = 2\delta_n ST(z)$, with $ST(z) = nST^x(z)$ and $ST^x(z)$ given in (17), and

$$\frac{[z^m]\widetilde{ST}(z)}{[z^m]P(z)} \sim 2\delta_n \frac{[z^m]ST(z)}{[z^m]P(z)} = 2\delta_n \mathbb{P}^{a,c}_m(True)$$
$$\sim 2\frac{(2\ln 2-1)}{2n}\frac{(2\ln 2-1)^2}{8n} = \frac{(2\ln 2-1)^3}{8n^2} + \mathcal{O}\left(\frac{1}{n^3}\right).$$

Moreover $g_x(z)$ is given by

$$g_x(z) = z + z\left(\exp\left(\sum_{\ell\geq 1}\frac{\hat{P}(z^\ell)-2z^\ell}{\ell}\right)-1\right),$$

and

$$g_x'(z) = 1 + \frac{1}{z}(g_x(z)-z) + g_x(z)\left(\sum_{\ell\geq 1} z^{\ell-1}(\hat{P}(z^\ell)-2)\right).$$

Since $g_x(\rho) \sim 2\rho$, we get

$$\lim_{z\to\delta_n}\frac{g_x'(z)}{P'(z)} \sim \lim_{z\to\delta_n}\frac{g_x(z)\hat{P}'(z)}{2\hat{P}'(z)-2n} \sim \frac{2\rho}{2} = \frac{2\ln 2-1}{2n},$$

and finally, with $\tilde{g}_x(z) = 2\delta_n g_x(z)$,

$$\frac{[z^m]\tilde{g}_x(z)}{[z^m]P(z)} \sim 2\delta_n\frac{[z^m]g_x(z)}{[z^m]P(z)} \sim 2\,\frac{(2\ln 2-1)}{2n}\,\frac{(2\ln 2-1)}{4n} = \frac{(2\ln 2-1)^2}{4n^2} + \mathcal{O}\left(\frac{1}{n^3}\right).$$ □

5. Limiting probability of a general function

In the previous sections, we have studied functions of complexity zero and one. In this section we are interested in the limiting probability of functions of higher complexity. To prove Theorem 2.5, Kozik showed that asymptotically almost all trees computing a function f have a "simple f" shape. To be more precise, they are obtained from a minimal tree by a single well-defined expansion, that is a special tree attached to a node of a minimal one. In this section we generalise this result to all models and give bounds for the number of such expansions.

5.1. The binary plane case.

The goal of this section is to prove existence and bound the constant λ_f appearing in Theorem 2.5. We show the following result.

Proposition 5.1. *For all Boolean functions f,*

$$\frac{8L(f)-3+\ell}{16^{L(f)}}M_f \leq \lambda_f \leq \frac{4L(f)^2+4L(f)-3}{16^{L(f)}}M_f$$

where M_f is the number of minimal trees representing f and

$$\ell = \begin{cases}\lceil\frac{L(f)}{2}\rceil & \text{for } L(f) > 1\\ 0 & \text{for } L(f) = 1.\end{cases}$$

The proof of this proposition is based on a result by Kozik [18]. He proved that the set of non negligible trees computing f is exactly the set of trees obtained by *expanding* a minimal tree of f once.

Remark: It is interesting to see that these bounds are equal when the complexity of the function is 1 and give the actual bound for literals we computed in Section 4.

Definition 5.2. Let t be an And/Or tree computing f, ν one of its nodes and t_ν the subtree rooted at ν. An *expansion* of t in ν is a tree obtained by replacing the subtree t_ν rooted at ν by a tree $t_\nu \diamond t_e$ where $\diamond \in \{\wedge, \vee\}$ and where t_e is an And/Or tree. We will say that such an expansion is *valid* when the expanded tree still computes f.

It follows from Kozik's results that the only non-negligible valid expansions that are to be considered are:

- The T-expansions: a valid expansion is a T-expansion if the inserted subtree t_e is a simple tautology (resp. a simple contradiction) and if the new label of ν is $\wedge$ (resp. $\vee$).
- The X-expansions: a valid expansion is an X-expansion if the inserted subtree t_e is (up to commutativity and associativity) of the shape $x \vee ...$ (resp. $x \wedge ...$) where x is an essential variable of f and if the new label of ν is $\wedge$ (resp. $\vee$).

We will elaborate this when discussing the analogous expansions for the associative case in Section 5.2.1.

In the following, we will call a T-expansion an $\wedge$-T-expansion (resp. an $\vee$-T-expansion) if the new label of ν is $\wedge$ (resp. $\vee$), and the same for X-expansions.

Proof of Proposition 5.1. In a Catalan And/Or tree, a T-expansion is possible in every node (without changing the computed function). At each node, we can expand by an $\vee$-T-expansion and by an $\wedge$-T-expansion, both on the right side and on the left side. As a minimal tree of f is of size $L(f)$, it has $2L(f) - 1$ nodes and there are $\lambda_T(f) = 4(2L(f) - 1)M_f$ different T-expansions that can be done in all minimal trees computing f.

We can now consider $\lambda_X(f)$, that is, the number of X-expansions (which do not change the computed function f). This number depends heavily on the shape of the minimal trees of f, therefore, we only give bounds for this number. An $\wedge$-X-expansion (resp. $\vee$-X-expansion) realized by x_i is valid at each leaf labelled by x_i, as well as at each node connected to one of them by an $\vee$-only (resp. $\wedge$-only) path, and at all of its sons. Let us note that

- At each leaf, we can do at least one $\vee$-X-expansion and one $\wedge$-X-expansion, both to the right and to the left, which gives a contribution of $4L(f)$. Further, we could do either one $\wedge$-X-expansion or one $\vee$-expansion to both sides at its father, depending on its level. But two different leaves having the same father could be labelled by the same variable. Hence this contributes $2\lceil L(f)/2 \rceil$ if $L(f) > 1$, as if $L(f) = 1$ a minimal tree consists of a single leaf with no father, but else all leaves share its father with one leaf of the same label in the worst case.
- at each node (internal or external), we can do at most 4 X-expansions (we choose between $\wedge$ and $\vee$ and between right and left side) for each different literal that appear on the leaves. There are at most $L(f)$ different literals appearing on the leaves of a minimal tree and a minimal tree has exactly $2L(f) - 1$ (internal or external) nodes. Therefore, $4L(f)(2L(f) - 1)M_f$ is an upper bound of $\lambda_X(f)$.

Therefore, we have the following bounds:

$$5L(f)M_f \leq \lambda_X(f) \leq 4L(f)(2L(f) - 1)M_f. \tag{18}$$

To end the proof of Proposition 5.1, we only need to note that

$$\frac{\lambda_f}{n^{L(f)+1}} = M_f \rho_n^{L(f)} \left(\lambda_T(f) w_1 + \lambda_X(f) w_2\right),$$

where w_1 is the limiting ratio of simple tautologies (resp. simple contradiction), and w_2 is the limiting ratio of trees of shape $x \vee ...$ for x a variable. Thanks to computations made in Section 3 (c.f. Theorem 3.6), we know that $w_1 = \frac{3}{4n}$. Moreover, the generating function g_x defined in

Section 3.1 counts exactly the number of trees that can be used for an X-expansion (according to a variable x). Therefore,

$$\lim_{z \to \rho_n} \frac{g_x'(z)}{T'(z)} \sim \frac{1}{2n} = w_2,$$

and with (18) we prove Proposition 5.1. □

5.2. The associative plane case.

The associative case appears to be similar to the binary plane case. We prove the following theorem:

Theorem 5.3. *In the associative plane model, let f be a non-constant Boolean function, whose complexity is denoted by $L(f)$. Then,*

$$\mathbb{P}_n^a(f) \sim \frac{\lambda_f^a}{n^{L(f)+1}},$$

when n tends to infinity, where λ_f^a is depending on the number of possible expansions of minimal trees of f. For $L(f) > 1$ we have

$$\left(\frac{3-2\sqrt{2}}{2}\right)^{L(f)} \left[133L(f) + 153 - (93L(f) + 108)\sqrt{2}\right] M_f \leq \lambda_f$$

$$\lambda_f \leq \left(\frac{3-2\sqrt{2}}{2}\right)^{L(f)} \left[-(12L(f)^2 - 247L(f) + 51) + (9L(f)^2 - 174L(f) + 36)\sqrt{2}\right] M_f,$$

where M_f is the number of minimal trees computing f.

To prove Theorem 5.3, we first have to prove that, as in the binary plane case, the set of non-negligible associative trees computing a Boolean function is the set of trees obtained from a minimal tree by expanding it once. Moreover, we have to find the non-negligible expansions that have to be considered. Then, we can prove Theorem 5.3 with the same methods as in the binary plane case.

5.2.1. Expansions of associative trees.

Because of the stratified structure of associative trees, we have to be careful with the definition of expansions, which is different to the one in the binary case:

Definition 5.4 (c.f. Figure 8). Let t be an And/Or associative tree computing f. We define two types of expansions of t.

- Let ν be an internal node of t (possibly the root) with subtrees $t_1, \dots, t_j$, $j \geq 2$. An expansion of the first kind of t in ν is a tree obtained by adding a subtree t_e to ν.
- Let ν be the root or a leaf of the tree. The tree obtained by replacing the subtree t_ν rooted at ν by $t_e \diamond t_\nu$, where $\diamond \in \{\wedge, \vee\}$ is chosen such that the obtained tree is stratified, is an expansion of t in ν of the second kind. In this case, $\diamond$ will be called the *new label* of ν.

We will say that such an expansion is *valid* when the expanded tree still computes f.

Proposition 5.5. *The set of non-negligible trees computing a Boolean function f is the set of trees obtained by expanding a minimal tree of f once. Moreover, the only non-negligible valid expansions are:*

- *The T-expansions: a valid expansion is a T-expansion if the inserted subtree t_e is a simple tautology (resp. a simple contradiction) and if the new label of ν is $\wedge$ (resp. $\vee$).*

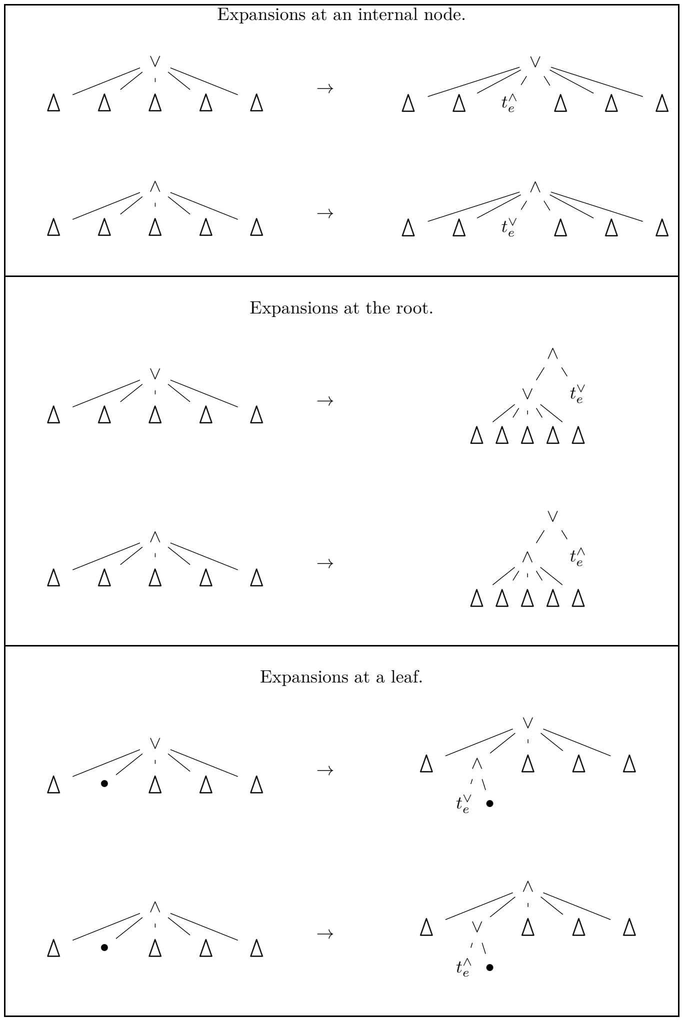


Figure 8: The different possible expansions in the associative case. Here, $t_e^{\vee}$ (resp. $t_e^{\wedge}$) stands for an associative tree rooted by $\vee$ (resp. $\wedge$). The tree pictured is only the subtree rooted at ν, before the expansion and after the expansion.

- *The X-expansions: a valid expansion is an X-expansion if the inserted subtree t_e is (up to commutativity) of the shape $x \vee \dots$ (resp. $x \wedge \dots$) where x is an essential variable of f and if the label of the father of t_e is $\wedge$ (resp. $\vee$).*

Before proving the result, let us introduce the pattern languages we will need:

$$\begin{cases} \hat{P} = \bullet | \check{P} \wedge \check{P} | \check{P} \wedge \check{P} \wedge \check{P} | \dots \\ \check{P} = \bullet | \hat{P} \vee \square | \hat{P} \vee \square \vee \square | \dots \\ S = \hat{P} | \check{P}; \\ \hat{N} = \bullet | \check{N} \wedge \square | \check{N} \wedge \square \wedge \square | \dots \\ \check{N} = \bullet | \hat{N} \vee \hat{N} | \hat{N} \vee \hat{N} \vee \hat{N} | \dots \\ R = \hat{N} | \check{N}. \end{cases} \tag{19}$$

Remark:

- The pattern language S has the following property: if all S-pattern leaves of a tree are set to $True$, then the whole tree itself computes $True$.

- The pattern language R has the following property: if all R-pattern leaves of a tree are set to $False$, then the whole tree itself computes $False$.

Definition 5.6. Let us consider the pattern languages L^r and $\bar{L} = L^{r+1}$. For $i \leq r$, a leaf is *on level* i if it is an $L^{(i)}$-pattern leaf but not an $L^{(i-1)}$-pattern leaf. An $\bar{L}$-pattern leaf which is not an L-pattern leaf is *on level* $r+2$.

Proof of Proposition 5.5. The proof of the corresponding result for binary trees, [18, Lemma 3.4]), can be taken almost verbatim. We only have to keep in mind that composition of pattern languages is restricted such that the resulting composed patterns are still stratified. But all the arguments still hold. For the sake of self-containedness we provide a sketch of the proof.

The idea is to take a tree computing f, and to replace every subtree which can be evaluated to $True$ or $False$ independently from the rest of the tree by a $\star$. Then the tree is simplified according to some rules (the rules stated in [18, Proof of Lemma 3.4] have an obvious analogue for associative trees) such that the tree contains no stars after all. Then, we state that the simplified tree t' is a minimal tree representing f.

Let f be a Boolean function whose complexity is $L(f)$ and let $\mathcal{V}$ denote the set of essential variables of f, We consider the pattern languages $L = R^{(L(f)+1)}[R \oplus S]$ and $\bar{L} = R^{(L(f)+1)}[(R \oplus S)^2]$, where the pattern leaves of $R \oplus S$ are all pattern leaves of the pattern R and the pattern S.

Let t be a tree of size $L(f)$ representing f. If the root of t is labelled with $\vee$ (resp. $\wedge$), then using a simple contradiction (resp. tautology) Φ, the new tree $\Phi \wedge t$ (resp. $\Phi \vee t$) still represents the function f. Since the limiting ratio of simple tautologies or contradictions is equal to $\Theta(1/n)$ and the $L(f)$ nodes of t are counted by $z^{L(f)}$, for sufficiently large n we obtain the lower bound $\mathbb{P}_n^a(f) \geq \frac{\alpha}{n^{L(f)+1}}$ and thus trees with at least $L(f)+2$ $(\bar{L}, \mathcal{V})$-restrictions can be neglected.

Further, it is possible to prove that trees with strictly less than $L(f)+1$ $(L, \mathcal{V})$-restrictions cannot represent f, because they would yield a tree of smaller size that $L(f)$ after the simplification process. Therefore all relevant trees have precisely $L(f)+1$ $(L, \mathcal{V})$- and as many $(\bar{L}, \mathcal{V})$-restrictions. But this implies that every variable appearing in a pattern leaf on level $L(f)+3$ is non essential and not repeated among the $\bar{L}$ pattern leaves and we may replace each subtree rooted on level $L(f)+3$ and having a parent node on level $L(f)+2$ by a star. Every leaf which is neither essential nor repeated in t will be replaced as well and one obtains a tree t^* which will be simplified according to some rules.

It turns out that the replacement of certain subtrees by wildcards and the subsequent simplification can be viewed as the reverse process of a single expansion of t'.

The second part of the proof is to understand which are the non-negligible valid expansions. Thanks to Theorem 3.10 and its generalizations, the trees obtained by expanding with a tree t_e with at least two $((R \oplus S)^2, \mathcal{V})$-restrictions are negligible. On the other hand there has to be at least one $((R \oplus S)^2, \mathcal{V})$-restriction in t_e, because if there was none, we could assign this tree to $False$ or $True$ independently from the rest of the tree. Since the expanded tree must still compute the function f, by simplification we would obtain a tree computing f being smaller than the minimal tree, which is impossible.

First case:. The tree t_e contains one repetition and no essential variable of f. Then, it has to compute a constant function (i.e. $True$ or $False$). If it does not, by previous arguments on tautologies, the subtree can be valuated to $True$ or $False$ independently from the rest of the tree. Thus, by simplification, we can obtain a tree, smaller than the minimal tree, computing f, which is a contradiction. Therefore, the expanding tree t_e is a simple tautology or a simple contradiction (thanks to Proposition 3.15). Moreover, as the expanded tree still has to compute f, if the father of t_e is an $\wedge$ (resp. $\vee$), t_e is a simple tautology (resp. contradiction), which gives a T-expansion.

Second case:. The subtree t_e contains no repetition and one essential variable, let us say x. Then, the essential variable has to appear on the first level. If it does not, the Boolean expression has shape $s_1 \wedge (s_2 \vee x)$ or $s_1 \vee (s_2 \wedge x)$ (up to commutativity). Moreover, the trees s_1 and s_2 have no $(R \oplus S, \mathcal{V})$-restrictions and therefore we can make them $False$ or $True$ independently from the rest of the tree. Then, we can valuate the whole tree either to $False$ or $True$ independently from x, which is impossible since x is an essential variable of f.

If an $\wedge$-X-expansion t_e according to the variable x_i is valid in a node ν, then every $\wedge$-X-expansion t'_e according to this variable x is valid at ν (and as well for $\vee$-X-expansions). □

5.2.2. Computing bounds for λ_f.

Proof of Theorem 5.3. As in the binary case, we have to compute the limiting ratio of T-expansions and X-expansions, and the number of nodes where the different kinds of expansions are allowed. Let us denote by M_f the number of minimal trees representing a given Boolean function f of complexity $L(f)$.

The limiting ratio of T-expansions is the limiting ratio of simple tautologies, which has already been computed in Section 3.2. We have that $w_1^a = \frac{51-36\sqrt{2}}{n}$.

Let g_x be the generating function of associative trees rooted at $\wedge$ (resp. $\vee$) and containing exactly one x in the first generation. Then,

$$g_x(z) = z\sum_{j\geq 2} j(A(z)-2z)^{j-1}.$$

Since the set of trees with more than one x in the first generation is negligible in front of the set of trees with exactly one x in the first generation, we can assume that

$$w_2^a = \lim_{z\to\alpha_n} \frac{g'_x(z)}{A'(z)} = \frac{3\sqrt{2}-4}{n}$$

is the limiting ratio of $\wedge$-X-expansions (resp. $\vee$-X-expansions).

As in the binary case, the number $\lambda_X(f)$ of X-expansions and the number $\lambda_T(f)$ of T-expansions allowed in a minimal tree depend on the shape of the considered minimal tree. Given a minimal tree t of f, let us number its internal nodes from 1 to N. Let us denote by $s(i)$ the number of sons of the internal node i. Moreover, let us denote by $d(i)$ the number of sons of the node i which are leaves. Then, if $\lambda_T(t)$ is the number of different T-expansions in the minimal tree t of f, we have that

$$\lambda_T(t) = 2L(f) + \sum_{i=1}^{N}(s(i)+1) + 2,$$

where $2L(f)$ is the number of different T-expansions allowed at the leaves of the tree (if the parent node is labelled by $\wedge$ (resp. $\vee$), only simple tautology (or contradiction respectively) T-expansions are allowed), $s(i)+1$ is the number of different T-expansions allowed at the node i (the number of different positions at node i is $s(i)+1$); and 2 is the number of expansions allowed at the root by pushing the root to the first generation and adding a new root with two sons. Therefore,

$$\lambda_T(t) = 2L(f) + \sum_{i=1}^{N} s(i) + N + 2 = 2L(f) + (L(f)+N-1) + N + 2,$$

and since $1 \leq N \leq L(f)-1$, we obtain that

$$3(L(f)+1)M_f \leq \lambda_T(f) \leq (5L(f)-1)M_f.$$

Further, given a leaf x_i, an $\wedge$-X-expansion realized by x_i is allowed at itself, at its father and at all its sisters (brothers that are reduced to a leaf), because two sisters cannot have the same label. Indeed, if two sisters have the same label (or even opposite labels), then, the considered tree can be simplified, and since we consider a minimal tree, this is impossible. Therefore, if $L(f) > 1$,

$$\lambda_X(t) = \sum_{i=1}^{N} d(i)(s(i)+1) + 2d(root) + 2\sum_{i=1}^{N} d(i)^2.$$

Lemma 5.7. *For all i, $d(i) \leq L(f) - N + 1$.*

Proof. Let us assume that there exist an internal node i_0 such that $d(i_0) > L(f) - N + 1$. It is easy to see that, as each node except the root has a unique father, $\sum_{i=1}^{N} d(i) = L(f) + N - 1$. Moreover,

$$\sum_{i=1}^{N} d(i) > \sum_{i \neq i_0} d(i) + (L(f) - N + 1) > 2(N-1) + (L(f) - N + 1)$$

since every internal node has at least two sons. Therefore, $\sum_{i=1}^{N} d(i) > L(f) + N - 1$, which is a contradiction. □

Therefore, thanks to the lemma,

$$\begin{aligned}
\lambda_X(t) &\leq \sum_{i=1}^{N} d(i) + (L(f) - N + 1) \sum_{i=1}^{N} s(i) + 2(L(f) - N + 1) + 2(L(f) - N + 1) \sum_{i=1}^{N} d(i) \\
&\leq L(f) + (L(f) - N + 1)[(N + L(f) - 1) + 2 + 2L(f)] \\
&\leq L(f) \cdot (3L(f) + 2).
\end{aligned}$$

On the other hand,

$$\lambda_X(t) \geq 3 \sum_{i=1}^{N} d(i) + 2 \sum_{i=1}^{N} d(i) = 5L(f),$$

and

$$\lambda_X(f) \geq 5L(f) M_f.$$

Finally, since

$$\frac{\lambda_f}{n^{L(f)+1}} = M_f \alpha_n^{L(f)} (\lambda_T(f) w_1^a + \lambda_X(f) w_2^a),$$

we get that

$$\left(\frac{3 - 2\sqrt{2}}{2}\right)^{L(f)} \left[133L(f) + 153 - (93L(f) + 108)\sqrt{2}\right] M_f \leq \lambda_f$$

$$\lambda_f \leq \left(\frac{3 - 2\sqrt{2}}{2}\right)^{L(f)} \left[-(12L(f)^2 - 247L(f) + 51) + (9L(f)^2 - 174L(f) + 36)\sqrt{2}\right] M_f. \quad □$$

5.3. The binary commutative case.

Theorem 5.8. *In the binary commutative case, let f be a non-constant Boolean function, whose complexity is denoted by $L(f)$. Then,*

$$\mathbb{P}_n^c(f) \sim \frac{\lambda_f^c}{n^{L(f)+1}},$$

when n tends to infinity, where λ_f^c is depending on the number of possible expansions of minimal trees of f, and

$$\frac{1794L(f) - 770}{512 \cdot 8^{L(f)}} M_f \leq \lambda_f^c \leq \frac{(2L(f) - 1)(512L(f) + 770)}{512 \cdot 8^{L(f)}} M_f,$$

where M_f is the number of minimal trees computing f.

Remark: It is interesting to see that these bounds are equal when the complexity of the function is 1 and give the limiting probability of literals computed in Section 4.

Proof. The proof relies completely on the binary plane case, doing minimal $[N]$-embeddings and $[N \oplus P]$-embeddings (the plane parts of an $[N \oplus P]$-embeddings are both the plane parts of an $[N]$-embedding or a $[P]$-embedding). It has been proven in Lemmas 3.25 and 3.26 that Theorem 3.19 can be applied to the pattern $[N \oplus P][\mathcal{C}]$, as the generating function of P is the same as the one of N. As in the proof of Theorem 3.17, embedding a tree $t \in \mathcal{C}$ into $N^{(L(f))}[N \oplus P]$ or $N^{(L(f))}[N \oplus P]^{(2)}$ represents an injection. Hence asymptotically, all trees computing a function f are obtained by a single expansion of a minimal tree of f. The calculation of the bounds can be done in the same way as in the plane binary case. If we denote by w_1^c the limiting ratio of simple tautologies and by w_2^c the limiting ratio of X-expansions. From Section 3.3 we know that $w_1^c = \frac{641}{1024n}$, and from Section 4 we know that $w_2^c = \frac{1}{2n}$. Moreover, since asymptotically all trees computing f are obtained by a single expansion of a minimal tree, we have $\frac{\lambda_f^c}{n^{L(f)+1}} = \gamma_n^{L(f)}(\lambda_T w_1^c + \lambda_X w_2^c)$ and

$$\lambda_T = 2(2L(f) - 1)M_f$$
$$2L(f)M_f \leq \lambda_X \leq 2L(f)(2L(f) - 1)M_f,$$

since $\gamma_n \sim \frac{1}{8n}$ when $n \to \infty$, the result follows. □

5.4. The associative commutative case.

Theorem 5.9. *In the associative and commutative model, let f be a non-constant Boolean function, whose complexity is denoted by $L(f)$:*

$$\mathbb{P}_n^{a,c}(f) \sim \frac{\lambda_f^{a,c}}{n^{L(f)+1}},$$

when n tends to infinity, where $\lambda_f^{a,c}$ is depending on the number of possible expansions of minimal trees of f. For $L(f) > 1$

$$\left(\frac{2\ln 2 - 1}{2}\right)^{L(f)} \left(\left(\ln^2 2 - \frac{1}{4}\right) L(f) + \ln^2 2 - 2\ln 2 + \frac{1}{2}\right) M_f \leq \lambda_f^{a,c}$$

$$\lambda_f^{a,c} \leq \left(\frac{2\ln 2 - 1}{2}\right)^{L(f)} \frac{(2\ln 2 - 1)(L(f) + 1 + 4\ln 2)L(f)}{4} M_f,$$

where M_f is the number of minimal trees computing f.

Proof. The result is easily proven by using the pattern $R^{(L(f))}[R \oplus S]$ and applying arguments of Sections 5.2 and 5.3. Therefore, we have $\frac{\lambda_f^{a,c}}{n^{L(f)+1}} = \delta_n^{L(f)}(\lambda_T w_1^{a,c} + \lambda_X w_2^{a,c})$. The calculation of the bounds is similar to the computations done in the plane associative case. From Section 3.4 we know that $w_1^{a,c} = \frac{(2\ln 2 - 1)^2}{4n}$ and in Section 4 we obtained $w_2^{a,c} = \frac{2\ln 2 - 1}{4n}$. Moreover, $\delta_n \sim \frac{2\ln 2 - 1}{2n}$. We can show

$$(L(f) + 2)M_f \leq \lambda_T \leq 2L(f)M_f$$
$$2L(f)M_f \leq \lambda_X \leq (L(f)^2 + 3L(f))M_f,$$

where the lower bound holds only for $L(f) > 1$, and the theorem is proven. □

6. Summary of results and conclusion

Finally, we have gained a better understanding of the influence of associativity and commutativity on the behaviour of the limiting distribution on Boolean functions induced by their tree representation. Indeed, we have shown that associativity and commutativity do not change the order of $\mathbb{P}_n(f)$ when n tends to infinity, it is still of order $\Theta\left(n^{-(L(f)+1)}\right)$. Section 5 gives bounds for the constants for a general function and shows that in all the models the already observed paradigm holds: Almost every tree representing a given Boolean function is a minimal tree expanded once. So the results exhibit a qualitatively similar behaviour of all four models. However, from a quantitative point of view associativity clearly yields a strong bias of the distribution against functions of small complexity whereas the effect of commutativity is much weaker (see Table 1 for an overview of the different constants). An intuitive explanation might be that binary or non-binary is a strong structural difference whereas plane or non-plane is not, since a random binary plane tree does not have many symmetry nodes (see [1] for a study of the number of symmetry nodes in binary trees). Interestingly, commutativity has almost no effect for binary trees, but a considerable effect in presence of associativity.

	Catalan trees	Associative (non-binary) trees	Commutative (non plane) trees	General trees
$True$	$\frac{3}{4} = 0.75$	$51 - 36\sqrt{2} \approx 0.0883$	$\frac{385}{512} \approx 0.75195$	$\frac{(2\ln 2 - 1)^2}{4} \approx 0.0373$
x	$\frac{5}{16} = 0.3125$	$546 - 386\sqrt{2} \approx 0.114$	$\frac{641}{2048} \approx 0.312988$	$\frac{(2\ln 2 - 1)^2(2\ln 2 + 1)}{4} \approx 0.0890$

Table 1: The different constants λ such that $\mathbb{P}(True) \sim \frac{\lambda}{n}$ and $\mathbb{P}(x) \sim \frac{\lambda}{n^2}$ when n tends to infinity, depending on the studied model of trees.

Note that the simple x trees we defined in Section 4 are exactly those trees obtained by expanding once a tree consisting of a single leaf x. Hence the proof of Proposition 4.2 is immediate.

We should also note that the relation between probability and complexity of a Boolean function holds for a fixed function f. It is not valid uniformly on *all* Boolean functions. For instance, we do not know the limiting probability of a function like $x_1 \wedge \cdots \wedge x_n$ where the set of essential variables depends on n. Such knowledge is important if we examine our models with respect to the *Shannon effect*: If we choose a Boolean function on n variables *uniformly* at random, asymptotically almost surely the function has a complexity which is exponential in n. In our models, we are still unable to compute the average complexity of a Boolean function. Further work (similar to [15] on implicational logic) is required before proving or disproving the presence of the Shannon effect for these non-uniform probability distributions.

References

[1] M. Bóna and P. Flajolet. Isomorphism and symmetries in random phylogenetic trees. *Journal of Applied Probability*, 46(4):1005–1019, 2009.

[2] R. B. Boppana. Amplification of probabilistic Boolean formulas. In *Proceedings of the 26th IEEE Symposium on Foundations of Computer Science*, pages 20–29, 1985.

[3] A. Brodsky and N. Pippenger. The Boolean functions computed by random Boolean formulas or how to grow the right function. *Random Structures and Algorithms*, 27:490–519, 2005.

[4] B. Chauvin, P. Flajolet, D. Gardy, and B. Gittenberger. And/Or trees revisited. *Combinatorics, Probability and Computing*, 13(4-5):475–497, July-September 2004.

[5] B. Chauvin, D. Gardy, and C. Mailler. The growing tree distribution for Boolean functions. In *8th SIAM Workshop on Analytic and Combinatorics (ANALCO)*, pages 45–56, 2011.

[6] M. Drmota. *Random trees.* Springer, Vienna-New York, 2009.

[7] M. Dubiner and U. Zwick. Amplification by read-once formulas. *SIAM Journal on Computing*, 26(1):15–38, 1997.

[8] P. Flajolet and A. M. Odlyzko. Singularity analysis of generating functions. *In SIAM J. Discrete Math.*, 3:216–240, 1990.

[9] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge U.P., Cambridge, 2009.

[10] H. Fournier, D. Gardy, and A. Genitrini. Balanced And/Or trees and linear threshold functions. In *6th SIAM Workshop on Analytic and Combinatorics (ANALCO)*, pages 51–57, 2009.

[11] H. Fournier, D. Gardy, A. Genitrini, and B. Gittenberger. Complexity and limiting ratio of Boolean functions over implication. In *33rd International Symposium on Mathematical Foundations of Computer Science (MFCS'08)*, pages 347–362, Torun, Pologne, August 2008.

[12] H. Fournier, D. Gardy, A. Genitrini, and B. Gittenberger. The fraction of large random trees representing a given Boolean function in implicational logic. *Random Structures and Algorithms*, 40(3):317–349, 2012.

[13] H. Fournier, D. Gardy, A. Genitrini, and M. Zaionc. Classical and intuitionnistic logic are asymptotically identical. In Springer-Verlag, editor, *Annual Conference on Computer Science Logic (CSL'07)*, pages 177–193, Lausanne, Suisse, 2007.

[14] D. Gardy. Random Boolean expressions. In *Colloquium on Computational Logic and Applications*, volume AF, pages 1–36. DMTCS Proceedings, 2006.

[15] A. Genitrini and B. Gittenberger. No Shannon effect on probability distributions on Boolean functions induced by random expressions. In *21st International Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, Vienna, Austria, july 2010. DMTCS Proceedings.

[16] A. Genitrini, B. Gittenberger, V. Kraus, and C. Mailler. Probabilities of Boolean functions given by random implicational formulas. *Electronic Journal of Combinatorics*, 19(2):P37, 20 pages, (electronic), 2012.

[17] A. Gupta and S. Mahajan. Using amplification to compute majority with small majority gates. *Computational Complexity*, 6(1):46–63, 1997.

[18] J. Kozik. Subcritical pattern languages for And/Or trees. In *Fifth Colloquium on Mathematics and Computer Science*, Blaubeuren, Germany, september 2008. DMTCS Proceedings.

[19] H. Lefmann and P. Savický. Some typical properties of large And/Or Boolean formulas. *Random Structures and Algorithms*, 10:337–351, 1997.

[20] M. Moczurad, J. Tyszkiewicz, and M. Zaionc. Statistical properties of simple types. *Mathematical Structures in Computer Science*, 10(5):575–594, 2000.

[21] J. B. Paris, A. Vencovská, and G. M. Wilmers. A natural prior probability distribution derived from the propositional calculus. *Annals of Pure and Applied Logic*, 70:243–285, 1994.

[22] G. Pólya and R. C. Read. *Combinatorial enumeration of Groups, Graphs and Chemical Compounds*. Springer Verlag, New York, 1987.

[23] R. A. Servedio. Monotone Boolean formulas can approximate monotone linear threshold functions. *Discrete Applied Mathematics*, 142(1-3):181–187, 2004.

[24] L. Valiant. Short monotone formulae for the majority function. *Journal of Algorithms*, 5:363–366, 1984.

[25] A. Woods. On the probability of absolute truth for And/Or formulas. *Bulletin of Symbolic Logic*, 12(3), 2005.