



HAL
open science

Subcritical pattern languages for and/or trees

Jakub Kozik

► **To cite this version:**

Jakub Kozik. Subcritical pattern languages for and/or trees. Fifth Colloquium on Mathematics and Computer Science, 2008, Kiel, Germany. pp.437-448, 10.46298/dmtcs.3582 . hal-01194664

HAL Id: hal-01194664

<https://inria.hal.science/hal-01194664>

Submitted on 7 Sep 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Subcritical pattern languages for and/or trees [†]

Jakub Kozik

Theoretical Computer Science, Jagiellonian University
Gronostajowa 3, Kraków, Poland.

`jkozik@tcs.uj.edu.pl`

Let $P_k(f)$ denote the density of and/or trees defining a boolean function f within the set of and/or trees with fixed number of variables k . We prove that there exists constant B_f such that $P_k(f) \sim B_f \cdot k^{-L(f)-1}$ when $k \rightarrow \infty$, where $L(f)$ denote the complexity of f (i.e. the size of a minimal and/or tree defining f). This theorem has been conjectured by Danièle Gardy and Alan Woods together with its counterpart for distribution π defined by some critical Galton-Watson process. Methods presented in this paper can be also applied to prove the analogous property for π .

Keywords: And/Or trees, probability distribution for Boolean functions, tree enumeration

1 Introduction

And/or trees are representations of boolean functions built from \wedge , \vee and positive and negative literals. Let us fix the set of variables to $\{x_1, \dots, x_k\}$. Since every and/or tree using at most variables $\{x_1, \dots, x_k\}$ defines some boolean function $f : \{0, 1\}^k \rightarrow \{0, 1\}$, the uniform probability distribution on the and/or trees of size n induces some probability distribution on the set of all boolean functions of k variables. It has been shown by Lefmann and Savický in [8] that for every fixed number of variables k when n (the size) tends to infinity the induced distribution on boolean functions converge to some (positive) probability distribution which we denote by P_k . The characterization of this limiting probability for big k is still a challenging problem. The authors of [8] suggested to study the relationships between $P_k(f)$ and so called complexity of function f which is the size of a minimal and/or tree defining f . They gave some lower and upper bounds for $P_k(f)$ dependent on the complexity. Although the upper bound have been improved in [1] by Chauvin, Flajolet, Gardy and Gittenberger, the gap is still quite large. The authors of [1] also introduced other interesting distribution on boolean function usually denoted by π , which results from some critical Galton-Watson process. The relationships between both distributions are a subject of ongoing research. For the general survey on the research concerning and/or trees we refer to [4].

[†]Research described in this paper was partially supported by POLONIUM grant (*Quantitative research in logic and functional languages*, cooperation between Jagiellonian University of Krakow, L' École Normale Supérieure de Lyon and Université de Versailles Saint-Quentin, contract number 7087/R07/R08) and French government research grant for young scientists (program number 0185)

In the recent paper [5], Gardy and Woods focused on special functions (namely constant function and so called "read-once" functions). For such a function f they analysed behaviour of $P_k(f)$ when k tends to infinity (in this approach f is treated as a boolean function of countably many variables, even though it depends only on the finite number of them). They stated the following conjecture:

For a "read-once" function f with complexity r , there exist constants b_f and B_f such that $\pi_k(f) \sim_k b_f k^{-r}$ and $P_k(f) \sim_k B_f k^{-r-1}$ as $k \rightarrow \infty$.

The main result of this paper is a proof of a variant of the conjectured property for all non-constant (not only "read-once") functions, for distributions P_k . Due to the lack of space we prove only that $P_k(f) = \Theta(k^{-r-1})$. We do not consider distributions π_k . However, the full generality of the conjecture and the analogous results for the distribution π_k can be obtained using the similar technique. The main tools we use are "subcritical tree pattern languages" defined in the section 2. As a simple example of applications of this technique, we present (in the subsection 3.2) quite simple proof of the result announced by Woods in [9] that "most of tautologies" are variations of formulae like $x \vee \bar{x} \vee \phi$.

1.1 Preliminaries

We use $n^{\underline{k}}$ to denote the falling factorial i.e. $n^{\underline{k}} = n(n-1) \dots (n-k+1)$. For any set of finite objects A we denote by $A(n)$ the number of elements of A of size n . All the trees we use are binary, planar and rooted. A tree language is any set of trees.

Generating functions. In the first part of this paper we make an extensive use of generating functions and singularity analysis (see [2]).

A function $f(z) \in \mathbb{R}[[z]]$ has singularity of the square root type in $\rho \in \mathbb{C}$ if it has Puiseux expansion around ρ of the form $f(z) = \sum_{n \in \mathbb{N}} c_n (\rho - z)^{\frac{n}{2}}$ with $c_1 \neq 0$. We use the following technical lemma, which is a consequence of the Theorem VII.8 from [2].

Lemma 1.1 *Let $f(z), g(z)$ be algebraic generating functions. Suppose that both functions have unique dominating singularities of the square root type in the same point $\rho \in \mathbb{R}_+$. Then the limit $\lim_{n \rightarrow \infty} \frac{[z^n]f(z)}{[z^n]g(z)}$ exists and is positive.*

Stirling numbers. After [7] we use the notation $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ for the Stirling number of the second kind. The number $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ is the number of partitions of a set of size n into k nonempty classes (subsets). We are going to use the following property.

Observation 1.2 *For a fixed $m \in \mathbb{N}$ the function $\left\{ \begin{smallmatrix} d \\ d-m \end{smallmatrix} \right\}$ is a polynomial of d .*

Proof: Let $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}_{>1}$ denote the number of partitions of a set of size n into k classes of size at least 2 each. Therefore for any fixed number n we have exactly

$$\binom{d}{n} \left\{ \begin{smallmatrix} n \\ n-m \end{smallmatrix} \right\}_{>1} \tag{1}$$

partitions of a set of size d in which exactly n elements belong to non singleton classes. For a fixed n and m the last equation is easily seen to be polynomial in d .

For every partition of a set of size d into $d-m$ classes there are at most $2m$ elements which do not belong to singleton classes. It means that $\left\{ \begin{smallmatrix} d \\ d-m \end{smallmatrix} \right\}$ is a finite sum of polynomials like (1). \square

Terms. Let $\text{Var} = \{x_1, x_2, x_3, \dots\}$ be a countable set of variables. An and/or tree is a planar rooted, binary tree, with internal nodes labelled by connectors \wedge, \vee , and leaves labelled by variables from Var or their negations (denoted by \bar{x}_1). The set of negated variables is denoted by $\overline{\text{Var}}$. The elements of

$\text{Var} \cup \overline{\text{Var}}$ are called literals. Throughout this paper we use many different kinds of trees, therefore we prefer to call and/or trees terms. The set of all terms is denoted by \mathcal{F} .

For every $k \in \mathbb{N}_1$, let \mathcal{F}_k be the set of all terms in which all used variables belong to the set $\text{Var}_k = \{x_1, \dots, x_k\}$.

For a term φ , the structure of φ is a tree which is constructed from φ by changing the labelling of all its leaves so that each leaf is labelled by \bullet . The set of all structures is denoted by \mathcal{T} . Leaf labelling functions are all the functions of type $\{1, \dots, n\} \rightarrow \text{Var} \cup \overline{\text{Var}}$ for some $n \in \mathbb{N}_1$. The leaf labelling of φ is the function $f : \{1, \dots, n\} \rightarrow \text{Var} \cup \overline{\text{Var}}$ such that $f(i)$ is the label of the i -th (e.g. from the left) leaf of φ . In most cases we work with terms from \mathcal{F}_k for some fixed $k \in \mathbb{N}_1$. Then we consider only labellings of type $f : \{1, \dots, n\} \rightarrow \text{Var}_k \cup \overline{\text{Var}}_k$.

The size of a structure is the total number of its leaves. The size of a formula is the size of its structure. The size of a labelling function $f : \{1, \dots, n\} \rightarrow \text{Var} \cup \overline{\text{Var}}$ equals n .

Note, that there is a natural (size-preserving) bijection between terms, and structure-labelling pairs of the same size. Therefore for every $k, n \in \mathbb{N}_1$ we have $\mathcal{F}_k(n) = \mathcal{T}(n) \cdot (2k)^n$.

We denote by $t(z)$ the generating function for the set of structures \mathcal{T} . Simple computations show that $t(z) = \frac{1}{4}(1 - \sqrt{1 - 8z})$, the radius of convergence is $\rho = \frac{1}{8}$, $t(z)$ has unique square root type singularity in ρ and $t(\rho) = \frac{1}{4}$.

Density. Instead of the definition of the distribution P_k sketched in the Introduction, we use equivalent approach based on densities. For any set of terms $A \subset \mathcal{F}_k$ we consider the limit $\lim_{n \rightarrow \infty} \frac{A(n)}{\mathcal{F}_k(n)}$. If the limit exists, it is called a density of A in \mathcal{F}_k . For a fixed boolean function of k variables f let A_f denote the set of terms from \mathcal{F}_k which define function f . It is easy to see that $\lim_{n \rightarrow \infty} \frac{A_f(n)}{\mathcal{F}_k(n)} = P_k(f)$ and the existence of the limit is granted by the result of Lefmann and Savicky from [8]. However, we lose connection with probability distribution on boolean functions when we consider set of terms which are essential subsets of A_f , while we still can talk about densities of such sets.

2 Pattern languages

A pattern language is any language of (binary, planar, rooted) trees with internal nodes labelled by \wedge or \vee and leaves labelled by the elements of $\{\bullet, \square\}$. The leaves labelled by \square are called placeholders (as opposed to regular leaves). For any tree language T and a pattern language P we define a language $P[T]$ as the language of all trees which can be constructed by substituting all placeholders \square in some pattern from P by some trees from T . P is unambiguous if for every T , every tree from $P[T]$ can be constructed in only one way. The size of the pattern is the number of its leaves labelled with \bullet (i.e. the number of regular leaves). For $d, h \in \mathbb{N}$ we denote by $P(d, h)$ the number of elements of P having d regular leaves and h placeholders. For a pattern language P we use bivariate generating function $p(x, y)$ with x marking regular leaves and y marking placeholders (i. e. $p(x, y) = \sum_{d \in \mathbb{N}, h \in \mathbb{N}} x^d y^h P(d, h)$).

For unambiguous P , in every tree $t \in P[T]$ we can distinguish set of its nodes (internal or leaves) which correspond to the nodes of pattern used to construct t . We call these nodes P -pattern nodes or simply pattern nodes if the pattern language is clear from the context. E.g. for the pattern $\square \wedge x_1$ the term $(x_2 \vee \bar{x}_2) \wedge x_1$, constructed from the pattern by substituting the placeholder by $(x_2 \vee \bar{x}_2)$, contains two pattern nodes: the root and the rightmost leaf (labelled by x_1 .)

Observation 2.1 Let P be an unambiguous pattern language, and T be a tree language. Let $p(x, y)$ be the generating function for P , and $t(z)$ for T . Then, the generating function of the language $P[T]$ is

$p(z, t(z))$.

Definition 2.2 Let $t(z)$ be a generating function having unique dominating singularity of the square root type in $\rho \in \mathbb{R}_+$. We say that the function $p(x, y)$ is subcritical for $t(z)$ if $p(x, y)$ is analytic in some set $\{(x, y) : |x| \leq \rho + \varepsilon, |y| \leq t(\rho) + \varepsilon\}$ for some $\varepsilon \in \mathbb{R}_+$.

We say that an unambiguous pattern language P is subcritical for a tree language T if the generating function $t(z)$ of T has unique dominating singularity of the square root type and the generating function $p(x, y)$ of P is subcritical for $t(z)$.

Observation 2.3 Let $t(z)$ be a generating function having unique dominating singularity of the square root type in $\rho \in \mathbb{R}_+$, and $p(x, y)$ be subcritical for $t(z)$. Suppose that both functions $t(z), p(x, y)$ have nonnegative coefficients as a formal power series. If there exist $\nu \in \mathbb{N}, \eta \in \mathbb{N}_1$ for which $[x^\nu y^\eta]p(x, y) > 0$, then the function $p(z, t(z))$ has unique dominating singularity of the square root type in ρ . In the opposite case its radius of convergence is strictly greater than ρ .

Proof: The function $t(z)$ is analytically continuable to the set $D = \{z \in \mathbb{C} : |z| < \rho + \varepsilon\} \setminus [\rho, \infty)$, for some positive ε . Since the composition of analytic functions is analytic, by subcriticality of $p(x, y)$ we get that the function $p(z, t(z))$ is analytic in D (for small enough ε). This shows that $p(z, t(z))$ can not have singularities in the set $|z| \leq \rho$ except for the point ρ .

Suppose that we have $\nu \in \mathbb{N}, \eta \in \mathbb{N}_1$ for which $[x^\nu y^\eta]p(x, y) > 0$. Then, by nonnegativity we get

$$[z^n]p(z, t(z)) \geq [z^n]t(z)^\eta z^\nu, \quad (2)$$

which shows that the radius of convergence of $p(z, t(z))$ cannot be greater than ρ . It means that $p(z, t(z))$ has unique dominating singularity in ρ .

A composition with analytic function cannot increase the branching type of algebraic singularity, therefore the function $p(z, t(z))$ has algebraic singularity at ρ with branching type at most 2. By the fact that $p(z, y)$ is bounded in the neighbourhood of $(\rho, t(\rho))$, we get that $\lim_{z \rightarrow \rho^-} p(z, t(z))$ is finite. It means that $p(z, t(z))$ has in ρ Puiseux expansion of the form: $p(z, t(z)) = c_0 + \sum_{n \in \mathbb{N}_1} c_n (\rho - z)^{\frac{n}{2}}$. If $c_1 = 0$, then by the standard algebraic asymptotics (see Theorem VII.8 from [2]) we would obtain that $\lim_{n \rightarrow \infty} \frac{[z^n]p(z, t(z))}{[z^n]t(z)} = 0$, which is impossible according to 2. It means that $c_1 \neq 0$, which proves that $p(z, t(z))$ has singularity of the square root type at ρ .

Suppose now that there are no $\nu \in \mathbb{N}, \eta \in \mathbb{N}_1$ for which $[x^\nu y^\eta]p(x, y) > 0$. Then by nonnegativity we get $p(z, t(z)) = p(z, 0)$, but by subcriticality, the radius of convergence of this function is greater than $\rho + \varepsilon$. \square

If $p(x, y)$ is subcritical for $t(z)$ then from the complex analysis we know that all partial derivatives $\frac{\partial^m p(x, y)}{(\partial x)^m}$, for $m \in \mathbb{N}$ are subcritical for $t(z)$ as well. According to the observation above it means that, for every $m \in \mathbb{N}$ the function $\frac{\partial^m p(x, y)}{(\partial x)^m} \Big|_{(z, t(z))}$ has dominating singularity of the square root type at ρ . In such a case using the Lemma 1.1 we get

$$\lim_{n \in \mathbb{N}} \frac{[z^n] \left(\frac{\partial^m p(x, y)}{(\partial x)^m} \Big|_{(z, t(z))} \right)}{[z^n]t(z)} = c_m$$

for some $c_m \in \mathbb{R}$. Additionally we get that c_m is positive whenever there are $\nu \in \mathbb{N}, \eta \in \mathbb{N}_1$ for which $[x^\nu y^\eta] \frac{\partial^m p(x, y)}{(\partial x)^m} > 0$.

2.1 Restrictions

Repetitions. For every term φ and any distinguished set of its leaves D we say that φ has m -repetitions among the leaves from D , if m equals the difference between the number of leaves in D and the number of distinct variables which occurs in these leaves (e.g. there are 3 repetitions among all the leaves of the term $((x_2 \wedge x_1) \vee \bar{x}_1) \vee (x_2 \wedge x_1)$). We say for short that φ has m D -repetitions.

Essential variables. Let k be a fixed number of allowed variables. In the section 3.3 we consider some distinguished subset of Var_k denoted by EVar . The elements of EVar are called essential variables. Suppose that the set EVar is fixed, and has cardinality l .

Definition 2.4 For a term φ and distinguished set of its leaves D , we say that φ has m restrictions among the leaves from D , if m equals the sum of D -repetitions and the number of different essential variables which have occurrences among the leaves from D . We say that φ has m D -restrictions for short.

We can rephrase the definition above as follows: let Var_D denote the set of variables with occurrences in D and $\#_D(x)$ denote the number of occurrences of a variable x in D , then the number of D -restrictions in φ equals:

$$\sum_{x \in \text{Var}_D \setminus \text{EVar}} (\#_D(x) - 1) + \sum_{x \in \text{Var}_D \cap \text{EVar}} \#_D(x).$$

For a pattern language P and $\varphi \in P[T]$, we say that φ has m P -restrictions if it has m restrictions among the P -pattern leaves.

Let t be a structure with size n and let D be some set of distinguished leaves of t . We are interested in the number of possible leaf labellings of t so that we obtain a term from \mathcal{F}_k with m restrictions among the leaves from D . Let d be a cardinality of D . (l is the cardinality of distinguished set $\text{EVar} \subset \mathcal{F}_k$ of essential variables.)

For any $r \leq m$ we calculate the number of labellings which gives r D -repetitions and m D -restrictions. Every such labelling determines some partition of the set D into $d - r$ classes of leaves labelled by the same variable (no matter positively or negatively). Additionally it must use $m - r$ essential variables. The number of such labellings is:

$$\left\{ \begin{matrix} d \\ d-r \end{matrix} \right\} \cdot \left\{ \begin{matrix} l \\ m-r \end{matrix} \right\} \cdot (d-r)^{m-r} \cdot (k-l)^{d-r-(m-r)} \cdot k^{n-d} \cdot 2^n,$$

where the consecutive factors correspond to: number of partitions of D into $d - r$ classes, number of choices of essential variables to be used for the leaves from D , number of assignments of essential variables to the classes of elements of D , number of assignments of non-essential variables to the remaining classes of elements of D , number of assignment of variables to the leaves not belonging to D , number of distributions of negations among all the leaves of constructed term.

Therefore the total number of labellings which gives terms from \mathcal{F}_k with m restrictions among the leaves is:

$$(k-l)^{d-m} \cdot k^{n-d} \cdot 2^n \cdot \sum_{r=0, \dots, m} \left\{ \begin{matrix} d \\ d-r \end{matrix} \right\} \cdot \left\{ \begin{matrix} l \\ m-r \end{matrix} \right\} \cdot (d-r)^{m-r}. \quad (3)$$

In case, when there are no essential variables we get much simpler form: $k^{d-m} \cdot k^{n-d} \cdot 2^n \cdot \left\{ \begin{matrix} d \\ d-m \end{matrix} \right\}$.

The observation below is a straightforward consequence of the Observation 1.2.

Observation 2.5 For fixed $l, m \in \mathbb{N}_1$ a function

$$w_{l,m}(d) = \sum_{r=0, \dots, m} \binom{d}{d-r} \cdot \binom{l}{m-r} \cdot (d-r)^{m-r}$$

is a polynomial of d . It is easily seen to have nonnegative values for $d \in \mathbb{N}$.

In the next lemma we prove that for a big k the value $(k-l)^{\underline{d-m}}$ is well estimated by $k^{\underline{d-m}}$

Lemma 2.6

$$k^{\underline{d-m}} - (k-l)^{\underline{d-m}} \leq (d-m) \cdot l \cdot k^{\underline{d-m-1}} + \binom{d-m}{2} \cdot k^{\underline{d-m-1}}$$

Proof: $k^{\underline{d-m}}$ is a number of sequences of variables of length $d-m$, while $(k-l)^{\underline{d-m}}$ is a number of sequences of different non-essential variables. Every sequence belonging to the difference of these sets contains either some essential variable (the number of such sequences is smaller than $(d-m) \cdot l \cdot k^{\underline{d-m-1}}$) or some repeated variable (the number of such sequences is smaller than $\binom{d-m}{2} \cdot k^{\underline{d-m-1}}$). \square

Lemma 2.7 Let $T \subset \mathcal{T}$ be a tree language whose generating function $t(z)$ has unique dominating singularity in $\rho \in \mathbb{R}_+$ of the square root type. Let P be an unambiguous pattern language, which is subcritical for T . Let $P[T](n, d)$ denote the number of trees from $P[T]$ of size n containing exactly d pattern leaves, and $w(d)$ be a nonzero polynomial of degree γ . Then

$$\lim_{n \rightarrow \infty} \frac{\sum_{d \in \mathbb{N}} P[T](n, d) w(d)}{T(n)} = c_w.$$

for some nonnegative real c_w . If additionally $w(d)$ has nonnegative values for all elements of \mathbb{N} and there exists integer $r \geq \gamma$ for which $w(r) > 0$ and P contains a pattern with r regular leaves and at least one placeholder, then $c_w \neq 0$.

Proof: Let $\alpha_\gamma d^\gamma + \alpha_{\gamma-1} d^{\gamma-1} + \dots + \alpha_0$ be a representation of $w(d)$. For $p(x, y) = \sum_{d \in \mathbb{N}, h \in \mathbb{N}} x^d y^h P(d, h)$ being the generating function for P we have:

$$x^j \frac{\partial^j p(x, y)}{(\partial x)^j} = \sum_{d \in \mathbb{N}, h \in \mathbb{N}} x^d y^h P(d, h) d^{\underline{j}}$$

Therefore

$$\alpha_\gamma x^\gamma \frac{\partial^\gamma p(x, y)}{(\partial x)^\gamma} + \alpha_{\gamma-1} x^{\gamma-1} \frac{\partial^{\gamma-1} p(x, y)}{(\partial x)^{\gamma-1}} + \dots + \alpha_0 = \sum_{d \in \mathbb{N}, h \in \mathbb{N}} x^d y^h P(d, h) w(d),$$

we denote this function by $p_w(x, y)$. Since the function $p(x, y)$ was subcritical for $t(z)$ all its derivatives in x multiplied by polynomials are subcritical as well, and so is any finite combination of them. It means that $p_w(x, y)$ is subcritical for $t(z)$, hence by the Observation 2.3 the function $p_w(z) = p_w(x, y)|_{(z, t(z))}$ has unique dominating singularity of the square root type in ρ or its radius of convergence is greater than ρ . By the Lemma 1.1 we get $\lim_{n \rightarrow \infty} \frac{[z^n] p_w(z)}{[z^n] t(z)} = c_w$ for some $c_w \in [0, \infty)$ (if the radius of

convergence of $p_w(z)$ is strictly greater than ρ then $c_w = 0$). The function $p_w m(z)$ is exactly the generating function of the sequence $\sum_{d \in \mathbb{N}} P[T](n, d)w(d)$ and $t(z)$ of the sequence $T(n)$, therefore $\lim_{n \rightarrow \infty} \frac{\sum_{d \in \mathbb{N}} P[T](n, d)w(d)}{T(n)} = c_w$.

Suppose that there exists integer $r \geq \gamma$ for which $w(r) > 0$ and P contains a pattern with r regular leaves and $h > 0$ placeholders. By nonnegativity of values of $w(d)$, and coefficients of $P[T](n, d)$ for every $i, j \in \mathbb{N}$ we have $[x^i y^j] p_w(x, y) \geq [x^i y^j](x^r y^h w(r))$. The inequality is preserved after substitution of the pair of nonnegative series $(z, t(z))$, hence $[z^i] p_w(z) \geq [z^i](z^r t(z)^h w(r))$. It shows that $p_w(z)$ must have singularity in ρ , which by the Lemma 1.1 gives $c_w \neq 0$. \square

Lemma 2.8 *Let P be an unambiguous pattern language, which is subcritical for \mathcal{T} . We denote by $\mathcal{F}_k^{[m]}(P[\mathcal{T}])(n)$ (resp. by $\mathcal{F}_k^{[\geq m]}(P[\mathcal{T}])(n)$) the number of terms from \mathcal{F}_k of size n whose structure belongs to $P[\mathcal{T}]$ and which have m (resp. at least m) P -restrictions. For every $m \in \mathbb{N}$ for which P contains pattern with at least $m + 1$ regular leaves and at least one placeholder, we have*

$$\lim_{n \rightarrow \infty} \frac{\mathcal{F}_k^{[\geq m]}(P[\mathcal{T}])(n)}{\mathcal{F}_k(n)} \sim_k \lim_{n \rightarrow \infty} \frac{\mathcal{F}_k^{[m]}(P[\mathcal{T}])(n)}{\mathcal{F}_k(n)} \sim_k \frac{c_{m,l}}{k^m}, \quad (4)$$

for some $c_{m,l} \in \mathbb{R}_+$ (the constant $c_{m,l}$ depends on the number of essential variables l).

Proof: Let $P[\mathcal{T}](n, d)$ denote the number of trees from $P[\mathcal{T}]$ of size n containing exactly d pattern leaves. For any fixed d and number of variables k we have $P[\mathcal{T}](n, d) \cdot w_{m,l}(d) \cdot k^{d-m} \cdot k^{n-d} \cdot 2^n$ terms from \mathcal{F}_k of size n with d pattern leaves and m restrictions among them. Then we have

$$\frac{\mathcal{F}_k^{[m]}(P[\mathcal{T}])(n)}{\mathcal{F}_k(n)} = \frac{2^n \cdot \sum_{d \in \mathbb{N}} (P[\mathcal{T}](n, d) \cdot w_{m,l}(d) \cdot (k-l)^{d-m} \cdot k^{n-d})}{T(n) \cdot (2k)^n}.$$

It implies

$$\frac{\mathcal{F}_k^{[m]}(P[\mathcal{T}])(n)}{\mathcal{F}_k(n)} \leq \frac{\sum_{d \in \mathbb{N}} (P[\mathcal{T}](n, d) \cdot w_{m,l}(d) \cdot k^{d-m} \cdot k^{n-d})}{T(n) \cdot k^n} \quad (5)$$

and by the Lemma 2.6

$$\frac{\mathcal{F}_k^{[m]}(P[\mathcal{T}])(n)}{\mathcal{F}_k(n)} \geq \frac{\sum_{d \in \mathbb{N}} P[\mathcal{T}](n, d) \cdot w_{m,l}(d) \cdot (k^{d-m} - k^{d-m-1}((d-m)l + \binom{d-m}{2})) \cdot k^{n-d}}{T(n) \cdot k^n}. \quad (6)$$

Applying the Lemma (2.7) we get:

$$\frac{\mathcal{F}_k(P[\mathcal{T}])^{[m]}(n)}{\mathcal{F}_k(n)} \leq \frac{\sum_{d \in \mathbb{N}} P[\mathcal{T}](n, d) \cdot w_{m,l}(d) \cdot k^{d-m} \cdot k^{n-d}}{T(n)k^n} \sim \frac{c_{m,l}}{k^m},$$

for some positive $c_{m,l}$ (it is easy to verify that $w_{m,l}(d)$ and P satisfy additional assumptions). Analogously we have

$$\frac{\sum_{d \in \mathbb{N}} P[\mathcal{T}](n, d) \cdot w_{m,l}(d) \cdot (l \cdot (d-m) + \binom{d-m}{2}) \cdot k^{n-m-1}}{T(n)k^n} = O\left(\frac{1}{k^{m+1}}\right),$$

as a result of application of the same lemma with the polynomial $w_{m,l}(d)(l \cdot (d - m) + \binom{d-m}{2})$. The last two estimations, together with 5 and 6 gives

$$\lim_{n \rightarrow \infty} \frac{\mathcal{F}_k^{[m]}(P[\mathcal{T}])(n)}{\mathcal{F}_k(n)} \sim_k \frac{c_{m,l}}{k^m}.$$

Finally, by the fact that the upper bound from (5) is greater than the fraction of all terms from \mathcal{F}_k containing at least m P -restrictions, we get

$$\lim_{n \rightarrow \infty} \frac{\mathcal{F}_k^{[m]}(P[\mathcal{T}])(n)}{\mathcal{F}_k(n)} \sim_k \lim_{n \rightarrow \infty} \frac{\mathcal{F}_k^{[\geq m]}(P[\mathcal{T}])(n)}{\mathcal{F}_k(n)}.$$

□

3 Applications to and/or trees

3.1 Pattern languages and subcriticality

Two most important pattern languages we use are defined as follows:

$$N = N \vee N | \square \wedge N | \bullet, \quad P = \square \vee P | P \wedge P | \bullet. \quad (7)$$

It is easy to observe that both are unambiguous for any tree language. For every term t if some valuation evaluates all its literals in N -pattern (resp. P -pattern) leaves to *False* (resp. *True*) then the whole term is evaluated to *False* (resp. *True*).

We consider also patterns which are compositions of patterns. For a pattern language S the pattern language $S^{(i)}$ is defined as $S^{(i-1)}[S]$ for $i > 1$ and as S for $i = 1$. Clearly, a composition of unambiguous pattern languages remains unambiguous. Let S_1, S_2 be unambiguous pattern languages with generating functions $s_1(x, y), s_2(x, y)$. It is easy to observe that the generating function for $S_1[S_2]$ equals $s_1(x, s_2(x, y))$.

Both pattern languages N and P have the same generating function $p(x, y)$ which satisfies the equation

$$p(x, y) = p(x, y)^2 + y \cdot p(x, y) + x.$$

Therefore we have

$$p(x, y) = \frac{1}{2}(-(y-1) - \sqrt{(y-1)^2 - 4x}).$$

(We discard the second solution using the fact that $p(x, 0)$ should be equal to the generating function of binary trees without any labels.)

The polynomial $(y-1)^2 - 4x$ has no zeros within the set $D = \{(x, y) \in \mathbb{C}^2 : |x| \leq \frac{1}{8} \wedge |y| \leq \frac{1}{4}\}$. It shows that $p(x, y)$ is analytic in this set. Finally, by the fact that $|t(z)| \leq \frac{1}{4}$ for $|z| \leq \frac{1}{8}$ we obtain subcriticality of P and N for \mathcal{T} . By nonnegativity of the coefficients of $p(x, y)$ as a formal power series we get

$$\max_{(x,y) \in D} |p(x, y)| = p\left(\frac{1}{8}, \frac{1}{4}\right) = \frac{1}{4}.$$

It shows that for an unambiguous pattern language S which is subcritical for \mathcal{T} , the pattern language $S[N]$ (resp. $S[P]$) is subcritical for \mathcal{T} as well.

We are going to use also combined pattern language $P \oplus N$, such that for any tree a leaf is $P \oplus N$ -pattern leaf if it is P -pattern leaf or N -pattern leaf. This pattern language is also unambiguous. In fact it is not hard to see that every $P \oplus N$ pattern leaf is $N[P]$ -pattern leaf. As a result, the pattern language $P \oplus N$ is subcritical for \mathcal{T} and for any unambiguous pattern language S which is subcritical for \mathcal{T} , the pattern language $S[P \oplus N]$ is subcritical for \mathcal{T} as well.

Corollary 3.1 *For every $i \in \mathbb{N}$ pattern language $N^{(i)}[P \oplus N]$ is unambiguous and subcritical for \mathcal{T} .*

3.2 Simple Tautologies

Within this section we assume that the set of essential variables is empty. A term (and/or tree) is a simple tautology if up to commutativity and associativity it is equal to $x \vee \bar{x} \vee \varphi$ for some variable x and term φ .

We present a simple proof of the theorem announced by Woods in [9], which we are going to use in the next section.

Theorem 3.2 *The density of tautologies among and/or trees with k variables asymptotically (with k) equals the density of simple tautologies.*

Proof: We use the pattern N defined by the equation (7). It is a matter of algebraic calculations to prove that the density of simple tautologies is asymptotically $\frac{d}{k}$ (the pattern $S = \square \wedge \square | S \vee S | \bullet$ might be helpful), for some positive $d \in \mathbb{R}$.

Every term which has no opposite literals among the N -pattern leaves, can be falsified by the valuation sending all these literals to false. We consider terms with exactly one $N[N]$ -restriction. Every simple tautology belongs to this set. An $N[N]$ -pattern leaf is called a first level leaf if it is also N -pattern leaf, otherwise it is called a second level pattern leaf. Suppose that t is a tautology with exactly one $N[N]$ -restriction (it must be repetition since there are no essential variables). In such a tautology the literals containing the occurrences of the repeated variable must be opposite. If at least one of the literals is on the second level, then we have no N -repetitions, and such a formula can be falsified. Therefore both opposite literals must be on the first level. If the t is not a simple tautology, then we have at least one node v labelled with \wedge above at least one of the $N[N]$ leaves containing the repeated variable. Let $t_1 \wedge t_2$ be the subtree of t rooted at v . Since we have only one $N[N]$ -repetition we can be sure that there are no opposite literals among the $N[N]$ -pattern leaves of t which do not belong to the subtree t_2 . Now it is easy to see that valuating all those literals to *False* falsifies the whole term. It shows that all tautologies among the terms with exactly one $N[N]$ -repetitions are simple.

The theorem follows from the fact that the density of terms with at least two $N[N]$ repetitions is asymptotically of the order k^{-2} (Lemma 2.8) and terms without $N[N]$ -repetitions can not be tautologies. \square

3.3 D. Gardy, A. Woods conjecture

Within this section we prove the variant of the Conjecture 1 from [5]. We address only the case of the distribution P_k . The proof for the distribution π_k is based on the same methods.

By a boolean function we mean a function $f : \{0, 1\}^{\text{Var}} \rightarrow \{0, 1\}$ (i.e. a function which transforms valuations of variables from Var to boolean values). We say that a function f depends on the variable x_i when there exist two valuations $v_1, v_2 \in \{0, 1\}^{\text{Var}}$, which differ only in the valuation of the variable x_i such that $f(v_1) \neq f(v_2)$. All the functions we consider depend on the finite number of variables (every

function defined by some term has this property). In that convention every distribution P_k can be seen as a discrete probability distribution with finite support on the set of all functions of type $\{0, 1\}^{\text{Var}} \rightarrow \{0, 1\}$.

The generalised version of D. Gardy, A. Woods conjecture from [5] for the distribution P_k :

Theorem 3.3 *Let f be a boolean non-constant function defined by some minimal and/or tree with size r , then there exists positive constant B_f such that*

$$P_k(f) \sim_k B_f k^{-r-1}$$

Due to the limited space, we present a proof of the simplified version of the above theorem (Lemma 3.4), which (in our opinion) captures the most important part of it.

Let us fix function f as in the statement of the theorem. Without loss of generality we may assume that f depends on variables x_1, \dots, x_j . We fix the set of essential variables to $\text{EVar} = \{x_1, \dots, x_j\}$. Note that every minimal tree defining f must have r restrictions among all its leaves.

Lemma 3.4

$$P_k(f) = \Theta(k^{-r-1}).$$

Proof: For r being a complexity of f we consider pattern language $R = N^{(r+1)}[P \oplus N]$. By the Corollary 3.1, pattern language R is unambiguous and subcritical for \mathcal{T} . For $i \leq r + 1$ we say that R -pattern leaf is on the level i if it is $N^{(i)}$ -pattern leaf, but not $N^{(i-1)}$ -pattern leaf (there are no N^0 pattern leaves). An R -pattern leaf is on the level $r + 2$ if it is not $N^{(r+1)}$ -pattern leaf.

Let t be a term of the size r defining f . Every term of the form $\varphi \wedge t$, for φ being a tautology, defines function f . Since the density of tautologies is asymptotically equal to $\frac{\bar{d}}{k}$ we get the trivial lower bound for $P_k(f)$ which holds for big enough k :

$$P_k(f) \geq \frac{\bar{d}}{k^{r+1}}, \quad (8)$$

for $\bar{d} \in \mathbb{R}_+$. From this bound we know that we can neglect all terms with at least $r + 2$ restrictions among R -pattern leaves (by the Lemma 2.8 the density of them is of the order k^{-r-2}).

Using subcriticality of R it is easy to see that the trees which does not have leaves on the level $r + 2$ does not contribute to the density $P_k(f)$. We assume that considered terms have at least one leaf on the level $r + 2$. We show that such a term must contain at least $r + 1$ restrictions among R -leaves to define function f .

Let t be a term defining f with at most r restrictions among R -leaves.

Let i be the smallest number such that term t contains the same number of $N^{(i-1)}$ -restrictions that of $N^{(i)}$ -restrictions. It is easy to observe that such i exists and is not greater than $r + 1$ (t must contain at least one restriction on the first level.) In particular there are no essential variables on the i -th level. We consider two cases:

First case: t has at most $r - 1$ restrictions among $N^{(i)}$ pattern leaves. Since there are no essential variables on the i -th level, we can replace these leaves with leaves labelled by constant *False* without changing the function defined by the term. But then, by the properties of the pattern N , all the nodes on the level i are valuated constantly to *False*, and hence we can trim the tree by replacing all subtrees whose roots are on the level i by leaves containing constant *False*. Such trimmed tree still calculates the

same function f . By the same token we can substitute $False$ for all non-essential variables occurring in the trimmed tree. Finally we can get rid of the constants using the rules:

$$False \wedge \varphi \equiv False, \quad False \vee \varphi \equiv \varphi.$$

The resulting tree contains no constant leaves, calculates the same function as t , and has at most $(r - 1)$ restrictions among all leaves. It means also that the resulting tree has size not greater than $r - 1$ (since all occurrences of variables are essential), which contradicts the fact that the complexity of f is r .

Second case: t has exactly r restrictions among $N^{(i)}$ pattern leaves. Then there are no restrictions among R -pattern leaves below the i -th level. It means that every variable which occurs in R -pattern leaf on the level i or below is not essential and has exactly one occurrence among R -pattern leaves. We use special leaf labelling symbol $*$ called wildcard, the occurrence of $*$ means that the leaf can be valuated independently of all essential variables and all variables occurring in the other R -pattern leaves.

Let us consider a node v which is on the level $r + 2$, but whose parent is on the level $r + 1$. By the construction of the pattern R , all the $(N \oplus P)$ -pattern leaves in the subtree rooted at v are R -pattern leaves. Since these leaves are on the level $r + 2$ the variables occurring in them are not repeated anywhere among R -leaves (they are also non essential). If we want to valuate the node v to $True$ (resp. $False$) it suffices to evaluate all its positive (negative) leaves to $True$ ($False$). It means that the node v can be evaluated to any value we want, independently of valuations of any R -pattern leaves, which are not below v . We substitute every such node v by a leaf labelled with wildcard. Whatever values we substitute for the wildcard, the computed function is still f . Note also that all the leaves in such constructed tree are R -pattern leaves. Similarly we can substitute wildcard $*$ for every leaf of such constructed tree which is labelled by a non-essential variable which is not repeated among other leaves. Finally, we eliminate wildcards by the following rules (and their symmetric variations):

$$* \vee * \equiv * \quad * \wedge * \equiv *$$

$$* \vee \varphi \equiv True \quad * \wedge \varphi \equiv False \quad (9)$$

$$False \wedge \varphi \equiv False \quad True \vee \varphi \equiv True$$

$$True \wedge \varphi \equiv \varphi \quad False \vee \varphi \equiv \varphi, \quad (10)$$

where φ denotes a term which does not contain any wildcard. The interpretation of rules is straightforward (in case of rules 9 since $*$ can be valuated to any value, we can always chose valuation which valuates $* \vee \varphi$ to $True$ independently of the valuation of any variables occurring in the R -pattern leaves of the tree). The tree \hat{t} obtained after such elimination still computes function f and has no wildcards. The tree before elimination contained r restrictions, at least one wildcard, and no constants. All the wildcards vanish during elimination process. It means that one of the rules from 9 must be used at least once. But use of such rule removes some subtree which originally contained at least one leaf not labelled with wildcard. Since such a leaf must contain either essential or repeated variable, the application of the rule reduces the number of restrictions in the tree, by at least one. Therefore the tree \hat{t} contains at most $r - 1$ restrictions which contradicts the fact that the complexity of f is r .

We got contradictions in both cases which shows that a tree having at least one leaf on the level $r + 2$ and defining the function f must have at least $r + 1$ repetitions among R -pattern leaves. This observation together with the Lemma 2.8 gives that for big enough k we have $P_k(f) \leq \frac{c}{k^{r+1}}$. Therefore, by the trivial lower bound (equation 8) we get: $P_k(f) = \Theta(\frac{1}{k^{r+1}})$. \square

Using similar methods we can prove the existence of the constants B_f from the theorem 3.3.

4 Concluding remarks

The method of subcritical pattern language seems to be suitable to be applied to a lot of problems concerning relative densities of tree languages. Beside the presented applications it can be used to prove the full version of D. Gardy, A. Woods conjecture for both of distributions (i.e. P_k and π_k , we omit these proofs for the lack of space). Moreover, lots of the results on relative densities of propositional logics can be proven anew within this framework (e.g. [3], [6]).

Presented results explains the behaviour of $P_k(f)$ for fixed function f when k tends to infinity. In such situation the complexity of f is eventually much smaller than the worst possible complexity (which is exponential in k). We believe that a deeper investigation of a typical structures of terms defining f with a careful treatment of the lower order terms in the estimation of $P_k(f)$ would give a new upper bound for that probability, which would be nontrivial for functions with small complexity. This would partially supplement the upper bound known so far (see [1]):

$$(1 + O(1/k)) \exp(-c \frac{r}{k^2}),$$

(where r is the complexity of f , and $c > 0$, see [1]), which is easily seen to be trivial for sufficiently small complexity r .

References

- [1] B. Chauvin, P. Flajolet, D. Gardy, and B. Gittenberger, *And/or trees revisited.*, *Combinatorics, Probability & Computing* **13** (2004).
- [2] Phillipe Flajolet and Robert Sedgewick, *Analytic combinatorics*, in preparation, preprint available at <http://algo.inria.fr/flajolet/Publications/book.pdf>, 2008.
- [3] Hervé Fournier, Danièle Gardy, Antoine Genitrini, and Marek Zaionc, *Classical and intuitionistic logic are asymptotically identical*, *CSL*, 2007, pp. 177–193.
- [4] Danièle Gardy, *Random boolean expressions*, *Computational Logic and Applications, CLA '05* (René David, Danièle Gardy, Pierre Lescanne, and Marek Zaionc, eds.), *DMTCS Proceedings*, vol. AF, *Discrete Mathematics and Theoretical Computer Science*, 2005, pp. 1–36.
- [5] Danièle Gardy and Alan Woods, *And/or tree probabilities of boolean functions*, *2005 International Conference on Analysis of Algorithms* (Conrado Martínez, ed.), *DMTCS Proceedings*, vol. AD, *Discrete Mathematics and Theoretical Computer Science*, 2005, pp. 139–146.
- [6] Antoine Genitrini, Jakub Kozik, and Marek Zaionc, *Intuitionistic vs classical tautologies, quantitative comparison*, *Types*, 2007, to appear.
- [7] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete mathematics: a foundation for computer science*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1989.
- [8] Hanno Lefmann and Petr Savický, *Some typical properties of large and/or boolean formulas*, *Random Struct. Algorithms* **10** (1997), no. 3, 337–351.
- [9] Alan Woods, *On the probability of absolute truth for and/or formulas*, *Annual Conference of The Australian Association for Logic* 2006., 2006.