arXiv:1903.00677v1 [math.CO] 2 Mar 2019

TREE SERIES AND PATTERN AVOIDANCE IN SYNTAX TREES

SAMUELE GIRAUDO

Abstract. A syntax tree is a planar rooted tree where internal nodes are labeled on a graded set of generators. There is a natural notion of occurrence of contiguous pattern in such trees. We describe a way, given a set of generators and a set of patterns, to enumerate the trees avoiding this last. The method is built around inclusion-exclusion formulas forming a system of equations on formal power series of trees, and composition operations of trees. This does not require particular conditions on the set of patterns to avoid. We connect this result to the theory of nonsymmetric operads. Syntax trees are the elements of such free structures, so that any operad can be seen as a quotient of a free operad. Moreover, in some cases, the elements of an operad can be seen as trees avoiding some patterns. Relying on this, we use operads as devices for enumeration: given a set of combinatorial objects we want enumerate, we endow it with the structure of an operad, understand it in term of trees and pattern avoidance, and use our method to count them. Several examples are provided.

Contents

Date: March 5, 2019.
2010 *Mathematics Subject Classification.* 05C05, 05A15, 32A05, 18D50.
Key words and phrases. Tree; Pattern avoidance; Enumeration; Formal power series; Operads.

Introduction

The general problem of counting objects is of primary importance in combinatorics. Several approaches exist for this purpose. Here, we focus on a strategy having an algebraic flavor consisting in endowing a set X of combinatorial objects with operations in order to form algebraic structures. The point is that the algebraic study of X (minimal generating sets, relations between generators, morphisms, *etc.*) leads to enumerative results. Operads [LV12, Mé15, Gir18] are very interesting algebraic structures in this context. They encode the notion of substitution of combinatorial objects into another ones. Moreover, formal power series on operads [Cha02, Cha08] or colored operads [Gir19] (that are generalizations of usual formal power series) offer new methods for enumerative questions. This work is intended to be an application of the theory of operads to combinatorics and enumeration. As main contribution, we provide a tool to express the Hilbert series (that is, the generating series of the sequence of the dimensions) of an operad $\mathcal{O}$ given one of its presentations by generators and relations (satisfying some properties). When $\mathcal{O}$ is an operad on combinatorial objects, this provides a description of the generating series of these objects. This is the consequence of the fact that some operads can be seen as operads of trees avoiding some patterns, and is related with the deeper notions of Koszul operads [GK94], Poincaré-Birkhoff-Witt bases for operads [Hof10], and Gröbner bases for operads [DK10].

Our main combinatorial result consists, given a set $\mathcal{P}$ of syntax trees (that are some labeled planar rooted trees, where labels are taken from a fixed alphabet), to obtain a system of equations expressing the formal sum of all the trees avoiding $\mathcal{P}$ (as connected components in the trees). The presented solution is built around an inclusion-exclusion formula and use simple grafting operations on trees. By considering the projection of this system to usual formal power series, this leads to a system of equations for the generating series of the trees avoiding $\mathcal{P}$. It is also possible to add formal parameters into these systems to enumerate the trees according to some statistics. Methods to enumerate trees that avoid some patterns have been already provided in [Row10] for the case of unlabeled binary trees, in [Par93] and [Lod05] for the case of patterns with two internal nodes, and in [KP15] for the general case. Our method differs from the latter one both in the approach and in the obtained systems of equations. Indeed, in the previous reference, the authors use combinatorics and enumerative properties to show algebraic properties on operads (while in the present work, we use operads to obtain combinatorial results and to count objects). Moreover, we obtain different systems of equations and we have less requirements about the sets $\mathcal{P}$ to avoid (they can be infinite or some of their trees can be factors of other ones).

This document is organized as follows. Section 1 contains elementary definitions about syntax trees and formal power series of trees. In Section 2, we state the main question of the paper about pattern avoidance in syntax trees and provide its main result (Theorem 2.2.4). Next, Section 3 is devoted to explain how to use nonsymmetric set-operads as devices for the enumeration of families of combinatorial objects. For this, the elementary definitions about operads are exposed, and a notion of refined Hilbert series of an operad depending on an orientation of one its presentations by generators and relations is provided. The document ends with Section 4 where examples of enumerations of some families of combinatorial objects are reviewed. We provide, by using several operad structures, the enumeration of bicolored Schröder trees, Schröder trees, binary trees, m-trees, noncrossing trees, Motzkin paths, and directed animals. The tools provided by this work highlight some (already known or not) statistics on these objects.

General notations and conventions. For any integers a and c, $[a, c]$ denotes the set $\{b \in \mathbb{N} : a \leqslant b \leqslant c\}$ and $[n]$, the set $[1, n]$. The cardinality of a finite set S is denoted by $\# S$. If u is a word, its length is denoted by $|u|$ and for any position $i \in [|u|]$, u_i is the ith letter of u.

1. Syntax trees and series

This section begins by setting elementary definitions about syntax trees, the main combinatorial objects of this work. Next, we present series on trees and some operations on them.

1.1. Syntax trees.

We set here elementary definitions and notations about graded sets, syntax trees, and composition operations on syntax trees.

1.1.1. *Graded sets and alphabets.* A *graded set* is a set $\mathfrak{G}$ admitting a decomposition as a disjoint union of the form

$$\mathfrak{G} := \bigsqcup_{n \geqslant 1} \mathfrak{G}(n). \tag{1.1.1}$$

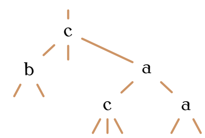
In the sequel, we shall call such a set an *alphabet* and each of its elements a *letter*. The *arity* $|x|$ of a letter x of $\mathfrak{G}$ is the unique integer n such that $x \in \mathfrak{G}(n)$. We say that $\mathfrak{G}$ is *combinatorial* if all the $\mathfrak{G}(n)$ are finite for all $n \geqslant 1$. In this case, the *generating series* of $\mathfrak{G}$ is the series $\mathcal{G}_{\mathfrak{G}}(t)$ defined by

$$\mathcal{G}_{\mathfrak{G}}(t) := \sum_{x \in \mathfrak{G}} t^{|x|}. \tag{1.1.2}$$

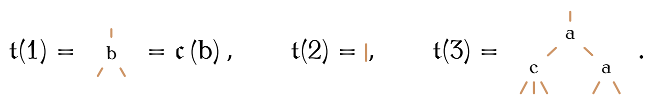
The coefficient of t^n in $\mathcal{G}_{\mathfrak{G}}(t)$ is $\# \mathfrak{G}(n)$ for any $n \geqslant 1$.

1.1.2. *Syntax trees.* Let $\mathfrak{G}$ be an alphabet. A *$\mathfrak{G}$-tree* (also called *$\mathfrak{G}$-syntax tree*) is a planar rooted tree such that its internal nodes of arity k are labeled by letters of arity k of $\mathfrak{G}$. Unless otherwise specified, we use in the sequel the standard terminology (such as *node*, *internal node*, *leaf*, *edge*, *root*, *child*, *etc.*) about planar rooted trees [Knu97] (see also [Gir18]). Let us set here some definitions about $\mathfrak{G}$-trees. The *degree* $\deg(\mathfrak{t})$ (resp. *arity* $|\mathfrak{t}|$) of a $\mathfrak{G}$-tree $\mathfrak{t}$ is its number of internal nodes (resp. leaves). The only $\mathfrak{G}$-tree of degree 0 and arity 1 is the *leaf* and is denoted by $|$. For any $\mathsf{a} \in \mathfrak{G}(k)$, the *corolla* labeled by a is the tree $\mathfrak{c}(\mathsf{a})$ consisting in one internal node labeled by a having as children k leaves. Given an internal node u of $\mathfrak{t}$, due to the planarity of $\mathfrak{t}$, the children of u are totally ordered from left to right and are thus indexed from 1 to the arity k of u. By assuming that the arity of the root of $\mathfrak{t}$ is k, for any $i \in [k]$, the *ith subtree* of $\mathfrak{t}$ is the tree $\mathfrak{t}(i)$ rooted at the ith child of $\mathfrak{t}$. Similarly, the leaves of $\mathfrak{t}$ are totally ordered from left to right and thus are indexed from 1 to $|\mathfrak{t}|$. The *height* of $\mathfrak{t}$ is the number of internal nodes belonging to a longest path connecting the root of $\mathfrak{t}$ to one of its leaves.

For instance, if $\mathfrak{G} := \mathfrak{G}(2) \sqcup \mathfrak{G}(3)$ with $\mathfrak{G}(2) := \{\mathsf{a}, \mathsf{b}\}$ and $\mathfrak{G}(3) := \{\mathsf{c}\}$,

$\mathfrak{t} :=$  (1.1.3)

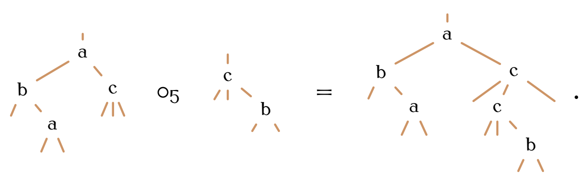
is a $\mathfrak{G}$-tree of degree 5, arity 8, and height 3. Its root is labeled by c and has arity 3. Moreover, we have

 (1.1.4)

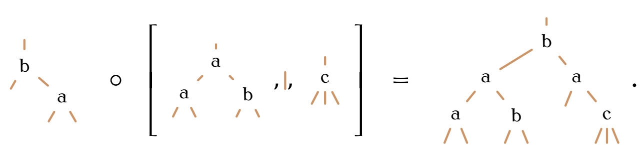
Given an alphabet $\mathfrak{G}$, we denote by $\mathbf{S}(\mathfrak{G})$ the graded set of all the $\mathfrak{G}$-trees where $\mathbf{S}(\mathfrak{G})(n)$ is the subset of $\mathbf{S}(\mathfrak{G})$ restrained on the $\mathfrak{G}$-trees of arity n. Observe that when $\mathfrak{G}$ is combinatorial and $\mathfrak{G}(1) = \emptyset$, $\mathbf{S}(\mathfrak{G})$ is combinatorial. In this case, the generating series $\mathcal{G}_{\mathbf{S}(\mathfrak{G})}(t)$ of $\mathbf{S}(\mathfrak{G})$, counting its elements with respect to their arities, satisfies

$$\mathcal{G}_{\mathbf{S}(\mathfrak{G})}(t) = t + \mathcal{G}_{\mathfrak{G}}\left(\mathcal{G}_{\mathbf{S}(\mathfrak{G})}(t)\right). \tag{1.1.5}$$

1.1.3. *Compositions of syntax trees.* Given $\mathfrak{t}, \mathfrak{s} \in \mathbf{S}(\mathfrak{G})$ and $i \in [|\mathfrak{t}|]$, the the *partial composition* $\mathfrak{t} \circ_i \mathfrak{s}$ is the $\mathfrak{G}$-tree obtained by grafting the root of $\mathfrak{s}$ onto the ith leaf of $\mathfrak{t}$. For instance, by considering the previous graded set $\mathfrak{G}$ of Section 1.1.2, one has

 (1.1.6)

Besides, given $\mathfrak{t} \in \mathbf{S}(\mathfrak{G})$ and $\mathfrak{s}_1, \dots, \mathfrak{s}_{|\mathfrak{t}|} \in \mathbf{S}(\mathfrak{G})$, the *full composition* $\mathfrak{t} \circ \left[\mathfrak{s}_1, \dots, \mathfrak{s}_{|\mathfrak{t}|}\right]$ is the $\mathfrak{G}$-tree obtained by grafting $\mathfrak{s}_i$ onto the ith leaf of $\mathfrak{t}$, simultaneously for all the $i \in [|\mathfrak{t}|]$. For instance, by considering the previous graded set $\mathfrak{G}$, one has

 (1.1.7)

By a slight but convenient abuse of notation, we shall in some cases simply write $\mathsf{a} \circ_i \mathsf{b}$ instead of $\mathfrak{c}(\mathsf{a}) \circ_i \mathfrak{c}(\mathsf{b})$, and write $\mathsf{a} \circ \left[\mathfrak{s}_1, \dots, \mathfrak{s}_{|\mathsf{a}|}\right]$ instead of $\mathfrak{c}(\mathsf{a}) \circ \left[\mathfrak{s}_1, \dots, \mathfrak{s}_{|\mathsf{a}|}\right]$ where a and b are letters of $\mathfrak{G}$ and $\mathfrak{s}_1$, ..., $\mathfrak{s}_{|\mathsf{a}|}$ are $\mathfrak{G}$-trees. Moreover, when the context is clear, we shall even write a for $\mathfrak{c}(\mathsf{a})$.

1.2. **Series on combinatorial sets.** We set here elementary definitions and notations about formal power series on arbitrary sets and about series on trees.

1.2.1. *Series on a set.* Let $\mathbb{K}$ be any field of characteristic zero. It is convenient, for enumerative purposes, to consider that $\mathbb{K}$ is simply the field $\mathbb{Q}$.

If X is a set, the linear span of X is denoted by $\mathbb{K}\langle X\rangle$. The dual space of $\mathbb{K}\langle X\rangle$, denoted by $\mathbb{K}\langle\langle X\rangle\rangle$ is by definition the space of the maps $\mathbf{f} : X \to \mathbb{K}$, called *$X$-series*. Let $\mathbf{f} \in \mathbb{K}\langle\langle X\rangle\rangle$. The coefficient $\mathbf{f}(x)$ of any $x \in X$ in $\mathbf{f}$ is denoted by $\langle x, \mathbf{f}\rangle$. The *support* of $\mathbf{f}$ is the set $\operatorname{Supp}(\mathbf{f}) := \{x \in X : \langle x, \mathbf{f}\rangle \neq 0\}$. We say that $x \in X$ *appears* in $\mathbf{f}$ if $x \in \operatorname{Supp}(\mathbf{f})$. By exploiting the vector space structure of $\mathbb{K}\langle\langle X\rangle\rangle$, any X-series $\mathbf{f}$ expresses as

$$\mathbf{f} = \sum_{x \in X} \langle x, \mathbf{f}\rangle\, x. \tag{1.2.1}$$

This notation using potentially infinite sums of elements of X accompanied with coefficients of $\mathbb{K}$ is common in the context of formal power series. In the sequel, we shall define and handle some X-series using the notation (1.2.1).

If $\mathbb{P}$ is a predicate on X, that is, for any $x \in X$, either $\mathbb{P}(x)$ holds or $\mathbb{P}(x)$ does not hold, the *predicate series* of $\mathbb{P}$ is the series

$$\operatorname{pr}(\mathbb{P}) := \sum_{\substack{x \in X \\ \mathbb{P}(x)}} x. \tag{1.2.2}$$

Moreover, for any subset Y of X, the *characteristic series* $\operatorname{ch}(Y)$ of Y is the predicate series of $\mathbb{P}$ where $\mathbb{P}(y)$ holds if and only if $y \in Y$. If $\mathbb{P}_1$ and $\mathbb{P}_2$ are two predicates on X, we denote by $\mathbb{P}_1 \wedge \mathbb{P}_2$ (resp. $\mathbb{P}_1 \vee \mathbb{P}_2$) the predicate wherein, for any $x \in X$, $(\mathbb{P}_1 \wedge \mathbb{P}_2)(x)$ (resp. $(\mathbb{P}_1 \vee \mathbb{P}_2)(x)$) holds if and only if $\mathbb{P}_1(x)$ and $\mathbb{P}_2(x)$ (resp. $\mathbb{P}_1(x)$ or $\mathbb{P}_2(x)$) hold.

Lemma 1.2.1. *Let X be a set and $\mathbb{P}_1, \dots, \mathbb{P}_n$, $n \geqslant 1$, be predicates on X. In $\mathbb{K}\langle\langle X\rangle\rangle$, we have*

$$\operatorname{pr}\left(\bigvee_{i \in [n]} \mathbb{P}_i\right) = \sum_{\substack{\ell \geqslant 1 \\ \{i_1, \dots, i_\ell\} \subseteq [n]}} (-1)^{1+\ell} \operatorname{pr}\left(\bigwedge_{j \in [\ell]} \mathbb{P}_{i_j}\right). \tag{1.2.3}$$

Proof. Let $\mathbf{f} := \operatorname{pr}(\mathbb{P}_1) + \operatorname{pr}(\mathbb{P}_2) - \operatorname{pr}(\mathbb{P}_1 \wedge \mathbb{P}_2)$ obtained from the right member of (1.2.3) in the particular case where $n = 2$. In $\mathbf{f}$, each $x \in X$ has a coefficient 0 or 1 according to the following rules:

1. if not $\mathbb{P}_1(x)$ and not $\mathbb{P}_2(x)$, then the coefficient of x is $0 + 0 - 0 = 0$;
2. if $\mathbb{P}_1(x)$ and not $\mathbb{P}_2(x)$, then the coefficient of x is $1 + 0 - 0 = 1$;
3. if not $\mathbb{P}_1(x)$ and $\mathbb{P}_2(x)$, then the coefficient of x is $0 + 1 - 0 = 1$;
4. if $\mathbb{P}_1(x)$ and $\mathbb{P}_2(x)$, then the coefficient of x is $1 + 1 - 1 = 1$.

Therefore, $\mathbf{f}$ is the series $\operatorname{pr}(\mathbb{P}_1 \vee \mathbb{P}_2)$, so that (1.2.3) holds for $n = 2$. Moreover, since (1.2.3) obviously holds when $n = 1$, by induction on n, the inclusion-exclusion formula of the statement of the lemma follows. □

1.2.2. *Series on syntax trees.* Let $\mathfrak{G}$ be an alphabet. We call *$\mathfrak{G}$-tree series* each series of $\mathbb{K}\langle\langle \mathbf{S}(\mathfrak{G})\rangle\rangle$. For any $n \geqslant 1$, the *composition product* of $\mathfrak{G}$-tree series is the product

$$\bar{\circ} : \mathbb{K}\langle\langle \mathbf{S}(\mathfrak{G})\rangle\rangle \otimes \mathbb{K}\langle\langle \mathbf{S}(\mathfrak{G})\rangle\rangle^{\otimes n} \to \mathbb{K}\langle\langle \mathbf{S}(\mathfrak{G})\rangle\rangle \tag{1.2.4}$$

defined for any $\mathfrak{G}$-tree series $\mathbf{f}$ and $\mathbf{g}_1, \dots, \mathbf{g}_n$ by

$$\mathbf{f}\,\bar{\circ}\,[\mathbf{g}_1, \dots, \mathbf{g}_n] := \sum_{\substack{\mathfrak{t} \in \mathbf{S}(\mathfrak{G})(n) \\ \mathfrak{s}_1, \dots, \mathfrak{s}_n \in \mathbf{S}(\mathfrak{G})}} \left(\langle \mathfrak{t}, \mathbf{f} \rangle \prod_{i \in [n]} \langle \mathfrak{s}_i, \mathbf{g}_i \rangle \right) \mathfrak{t} \circ [\mathfrak{s}_1, \dots, \mathfrak{s}_n]. \tag{1.2.5}$$

Observe that this product is linear in all its inputs, and that it can be seen as an extension by linearity of the full composition product of $\mathfrak{G}$-trees.

1.2.3. *Generating series.* Let us define from $\mathfrak{G}$ the set

$$\mathbf{Q}_{\mathfrak{G}} := \{q_{\mathtt{a}} : \mathtt{a} \in \mathfrak{G}\} \tag{1.2.6}$$

of formal parameters. The usual set of the commutative generating series on the set $\{t, q\} \cup \mathbf{Q}_{\mathfrak{G}}$ of parameters is denoted by $\mathbb{K}\langle\langle t, q, \mathbf{Q}_{\mathfrak{G}}\rangle\rangle$.

The *trace* $\mathrm{tr}(\mathfrak{t})$ of a $\mathfrak{G}$-tree $\mathfrak{t}$ is the monomial of $\mathbb{K}\langle\langle t, q, \mathbf{Q}_{\mathfrak{G}}\rangle\rangle$ defined by

$$\mathrm{tr}(\mathfrak{t}) := \prod_{\mathtt{a} \in \mathfrak{G}} q_{\mathtt{a}}^{\deg_{\mathtt{a}}(\mathfrak{t})}, \tag{1.2.7}$$

where for any $\mathtt{a} \in \mathfrak{G}$, $\deg_{\mathtt{a}}(\mathfrak{t})$ is the number of internal nodes of $\mathfrak{t}$ labeled by $\mathtt{a}$. Moreover, the *enumeration map* on $\mathbb{K}\langle\langle \mathbf{S}(\mathfrak{G})\rangle\rangle$ is the map

$$\mathrm{en} : \mathbb{K}\langle\langle \mathbf{S}(\mathfrak{G})\rangle\rangle \to \mathbb{K}\langle\langle t, q, \mathbf{Q}\rangle\rangle \tag{1.2.8}$$

defined linearly by

$$\mathrm{en}(\mathfrak{t}) := t^{|\mathfrak{t}|} q^{\deg(\mathfrak{t})}\, \mathrm{tr}(\mathfrak{t}). \tag{1.2.9}$$

For any $\mathfrak{G}$-tree series $\mathbf{f}$, the *enumerative image* of $\mathbf{f}$ is the generating series $\mathrm{en}(\mathbf{f})$. By definition, the coefficient of $t^n q^d q_{\mathtt{a}_1}^{\alpha_1} \dots q_{\mathtt{a}_\ell}^{\alpha_\ell}$, $n \geqslant 1$, $d \geqslant 0$, $\alpha_i \geqslant 0$, $i \in [\ell]$, in the enumerative image of the characteristic series of a set S of $\mathfrak{G}$-trees is the number of trees $\mathfrak{t}$ of S having n as arity, d as degree, and $q_{\mathtt{a}_1}^{\alpha_1} \dots q_{\mathtt{a}_\ell}^{\alpha_\ell}$ as trace.

Observe that for any alphabet $\mathfrak{G}$, since there are finitely many $\mathfrak{G}$-trees having a fixed trace, the enumerative image of any $\mathfrak{G}$-tree series is always well-defined. Moreover, when $\mathfrak{G}$ is combinatorial and $\mathfrak{G}(1) = \emptyset$, there are finitely many $\mathfrak{G}$-trees having a given arity $n \geqslant 1$. For this reason, for any set S of $\mathfrak{G}$-trees, the specialization $\mathrm{ch}\,(S)_{|q:=1, q_{\mathtt{a}}:=1, \mathtt{a} \in \mathfrak{G}}$ is well-defined and is the series wherein the coefficient of t^n is the number of $\mathfrak{G}$-trees of S of arity n. Observe finally that when $\mathfrak{G}$ is finite, there are finitely many $\mathfrak{G}$-trees having a given degree $d \geqslant 0$. For this reason, the specialization $\mathrm{ch}\,(S)_{|t:=1, q_{\mathtt{a}}:=1, \mathtt{a} \in \mathfrak{G}}$ is well-defined and is the series wherein the coefficient of q^d is the number of $\mathfrak{G}$-trees of S of degree d.

Proposition 1.2.2. *For any alphabet $\mathfrak{G}$, any $\mathfrak{G}$-tree $\mathfrak{t}$ of arity $n \geqslant 1$, and any $\mathfrak{G}$-tree series $\mathbf{f}_1, \dots, \mathbf{f}_n$,*

$$\mathrm{en}\left(\mathfrak{t}\,\bar{\circ}\,[\mathbf{f}_1, \dots, \mathbf{f}_n]\right) = \frac{1}{t^{|\mathfrak{t}|}} \mathrm{en}(\mathfrak{t}) \prod_{i \in [n]} \mathrm{en}\left(\mathbf{f}_i\right). \tag{1.2.10}$$

Proof. The statement of the proposition follows by computing the enumerative image of the right member of (1.2.5). □

Proposition 1.2.2 admits the following practical consequence. Assume that we have a set S of $\mathfrak{G}$-trees we want enumerate (with respect to the arities and the traces of its elements). A way to accomplish this consists in providing an expression for $\mathrm{en}(\mathrm{ch}(S))$. In the case where we have a description of $\mathrm{ch}(S)$ as an expression using the sum, the multiplication by a scalar, and the composition product of $\mathfrak{G}$-tree series, we obtain thanks to Proposition 1.2.2 an expression for $\mathrm{en}(\mathrm{ch}(S))$ using only the sum, the multiplication by a scalar, and the multiplication product of generating series. We shall use this observation in the sequel to obtain systems of equations of generating series from systems of equations of tree series.

2. Tree series and pattern avoidance

This section deals with two notions pattern avoidance in syntax trees: factor-avoidance and prefix-avoidance. The aim is to describe a way to enumerate the syntax trees factor-avoiding a set of patterns. For this, we begin by introducing some technical tools. Then, we state our main result, provide some of its consequences, and finish by reviewing some examples.

2.1. **Patterns in syntax trees.**

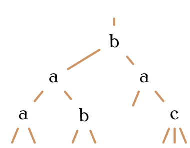
The notions of prefix, factor, and suffix in syntax trees are set here. Their immediate properties are stated.

2.1.1. *Factors, prefixes, and suffixes in trees.*

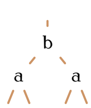
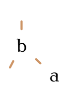
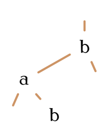
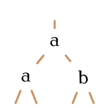
Let $\mathfrak{G}$ be an alphabet and let $\mathfrak{t}$ be a $\mathfrak{G}$-tree. When $\mathfrak{t}$ expresses as

$$\mathfrak{t} = \mathfrak{r} \circ_i \left(\mathfrak{s} \circ \left[\mathfrak{r}_1, \dots, \mathfrak{r}_{|\mathfrak{s}|}\right]\right) \tag{2.1.1}$$

for some $\mathfrak{G}$-trees $\mathfrak{s}$, $\mathfrak{r}$, and $\mathfrak{r}_1$, …, $\mathfrak{r}_{|\mathfrak{s}|}$, and $i \in [|\mathfrak{r}|]$, $\mathfrak{s}$ is a *factor* of $\mathfrak{t}$ and this property is denoted by $\mathfrak{s} \preccurlyeq_{\mathrm{f}} \mathfrak{t}$. Intuitively, this says that one can put down $\mathfrak{s}$ at a certain place into $\mathfrak{t}$, by possibly superimposing leaves of $\mathfrak{s}$ and internal nodes of $\mathfrak{t}$. When $\mathfrak{r} = |$ in (2.1.1), $\mathfrak{s}$ is a *prefix* of $\mathfrak{t}$ and this property is denoted by $\mathfrak{s} \preccurlyeq_{\mathrm{p}} \mathfrak{t}$. Intuitively, this says that $\mathfrak{s}$ is a factor of $\mathfrak{t}$ wherein the root of $\mathfrak{s}$ can be put down onto the root of $\mathfrak{t}$. Finally, when $\mathfrak{r}_j = |$ for all $j \in [|\mathfrak{s}|]$ in (2.1.1), $\mathfrak{s}$ is a *suffix* of $\mathfrak{t}$ and this property is denoted by $\mathfrak{s} \preccurlyeq_{\mathrm{s}} \mathfrak{t}$. Let us consider some examples. By setting

$\mathfrak{t} :=$  , (2.1.2)

we have

 $\preccurlyeq_{\mathrm{f}} \mathfrak{t}$,  $\preccurlyeq_{\mathrm{p}} \mathfrak{t}$,  $\preccurlyeq_{\mathrm{p}} \mathfrak{t}$,  $\preccurlyeq_{\mathrm{s}} \mathfrak{t}$. (2.1.3)

Proposition 2.1.1. *For any alphabet $\mathfrak{G}$, $\preccurlyeq_{\mathrm{f}}$, $\preccurlyeq_{\mathrm{p}}$, and $\preccurlyeq_{\mathrm{s}}$ endow $\mathbf{S}(\mathfrak{G})$ with poset structures. Moreover, the poset $(\mathbf{S}(\mathfrak{G}), \preccurlyeq_{\mathrm{f}})$ is an extension of $\left(\mathbf{S}(\mathfrak{G}), \preccurlyeq_{\mathrm{p}}\right)$.*

Proof. The fact that $\preccurlyeq_{\mathrm{f}}$, $\preccurlyeq_{\mathrm{p}}$, and $\preccurlyeq_{\mathrm{s}}$ are order relations is straightforward from their definitions. Moreover, since for any $\mathfrak{G}$-trees $\mathfrak{s}$ and $\mathfrak{t}$, $\mathfrak{s} \preccurlyeq_{\mathrm{p}} \mathfrak{t}$ implies $\mathfrak{s} \preccurlyeq_{\mathrm{f}} \mathfrak{t}$, the second part of the statement of the proposition holds. □

When $\mathfrak{s}$ is not a factor (resp. a prefix) of $\mathfrak{t}$, $\mathfrak{t}$ *factor-avoids* (resp. *prefix-avoids*) $\mathfrak{s}$. This property is denoted by $\mathfrak{s} \not\preccurlyeq_{\mathrm{f}} \mathfrak{t}$ (resp. $\mathfrak{s} \not\preccurlyeq_{\mathrm{p}} \mathfrak{t}$). By extension, when $\mathcal{P}$ is any subset of $\mathbf{S}(\mathfrak{G})$, $\mathfrak{t}$ *factor-avoids* (resp. *prefix-avoids*) $\mathcal{P}$ if for all $\mathfrak{s} \in \mathcal{P}$, $\mathfrak{s} \not\preccurlyeq_{\mathrm{f}} \mathfrak{t}$ (resp. $\mathfrak{s} \not\preccurlyeq_{\mathrm{p}} \mathfrak{t}$). By a slight abuse of notation, this property is denoted by $\mathcal{P} \not\preccurlyeq_{\mathrm{f}} \mathfrak{t}$ (resp. $\mathcal{P} \not\preccurlyeq_{\mathrm{p}} \mathfrak{t}$).

Lemma 2.1.2. *Let $\mathfrak{G}$ be an alphabet, and $\mathfrak{s}$ and $\mathfrak{t}$ be two $\mathfrak{G}$-trees. Then, $\mathfrak{s}$ is a prefix of $\mathfrak{t}$ if and only if $\mathfrak{s} = |$ or there exists a letter $\mathrm{a} \in \mathfrak{G}(k)$ such that $\mathfrak{s} = \mathrm{a} \circ [\mathfrak{s}(1), \dots, \mathfrak{s}(k)]$, $\mathfrak{t} = \mathrm{a} \circ [\mathfrak{t}(1), \dots, \mathfrak{t}(k)]$, and for all $i \in [k]$, $\mathfrak{s}(i) \preccurlyeq_{\mathrm{p}} \mathfrak{t}(i)$.*

Proof. This follows directly from the definition of the relation $\preccurlyeq_{\mathrm{p}}$. □

2.1.2. *Tree series avoiding factors.* For any subset $\mathcal{P}$ of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$, let $\mathbb{P}_{\mathcal{P}}$ be the predicate on $\mathbf{S}(\mathfrak{G})$ wherein $\mathbb{P}_{\mathcal{P}}(\mathfrak{t})$ holds if and only $\mathcal{P} \not\preccurlyeq_{\mathrm{f}} \mathfrak{t}$. Let also $\mathbf{F}(\mathcal{P})$ be the $\mathfrak{G}$-tree series defined by

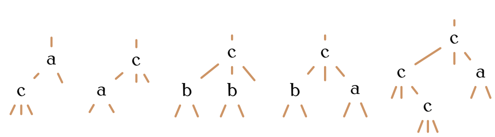
$$\mathbf{F}(\mathcal{P}) := \operatorname{pr}(\mathbb{P}_{\mathcal{P}}) . \tag{2.1.4}$$

In other terms, $\mathbf{F}(\mathcal{P})$ is the characteristic series of all $\mathfrak{G}$-trees factor-avoiding all trees of $\mathcal{P}$. In this context, we say that the elements of $\mathcal{P}$ are *patterns*. Notice that we consider only sets of patterns $\mathcal{P}$ such that $| \notin \mathcal{P}$ since there exists no $\mathfrak{G}$-tree factor-avoiding $|$. Notice also that, for the while, there is no restriction neither on $\mathfrak{G}$ nor or $\mathcal{P}$. This set $\mathcal{P}$ of patterns can be infinite or some trees can be themselves factors of another ones. The aim of the next section is to provide a system of equations to describe $\mathbf{F}(\mathcal{P})$ within the more general possible context.

2.2. **Pattern avoidance and enumeration.**

We provide here a way to obtain a system of equations to describe the $\mathfrak{G}$-tree series $\mathbf{F}(\mathcal{P})$. For this, we start by introducing tools, namely consistent words and admissible trees. From now, to not overload the notation, sets of patterns are denoted by omitting the braces and the commas. Hence, sets of patterns can be seen as unordered forests of $\mathfrak{G}$-trees without repeated trees.

Moreover, all examples of this section are based upon the finite set of patterns

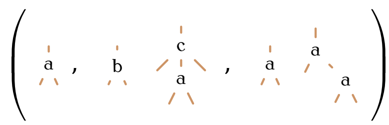
$\mathcal{P} :=$  . (2.2.1)

2.2.1. *Consistent words.* Let $\mathfrak{G}$ be an alphabet and $\mathcal{P}$ be a subset of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$. For any $\mathrm{a} \in \mathfrak{G}(k)$, $k \geqslant 1$, let

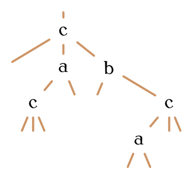
$$\mathcal{P}_{\mathrm{a}} := \{\mathfrak{s} \in \mathcal{P} : \mathfrak{c}(\mathrm{a}) \preccurlyeq_{\mathrm{p}} \mathfrak{s}\}. \tag{2.2.2}$$

In other words, $\mathcal{P}_{\mathrm{a}}$ is the subset of $\mathcal{P}$ of the patterns having roots labeled by a. A word $\mathcal{S} := (\mathcal{S}_1, \dots, \mathcal{S}_k)$ where each $\mathcal{S}_i$ is a subset of $\mathbf{S}(\mathfrak{G})$, $i \in [k]$, is *$\mathcal{P}_{\mathrm{a}}$-consistent* if for any $\mathfrak{s} \in \mathcal{P}_{\mathrm{a}}$, there is an $i \in [k]$ such that $\mathfrak{s}(i) \neq |$ and $\mathfrak{s}(i) \in \mathcal{S}_i$. Observe that when $\mathfrak{c}(\mathrm{a}) \in \mathcal{P}$, there is no $\mathcal{P}_{\mathrm{a}}$-consistent words. Moreover, a $\mathfrak{G}$-tree $\mathfrak{t}$ is *$\mathcal{S}$-admissible* if the root of $\mathfrak{t}$ is labeled by a and for all $i \in [k]$, $\mathfrak{t}(i)$ prefix-avoid $\mathcal{S}_i$.

For instance, by considering the set (2.2.1) of patterns, the word

$\mathcal{S} :=$  (2.2.3)

is $\mathcal{P}_{\mathrm{c}}$-consistent. Moreover, the tree

$\mathfrak{t} :=$  (2.2.4)

is $\mathcal{S}_{\mathrm{c}}$-admissible. Observe however that $\mathfrak{t}$ does not factor-avoids $\mathcal{P}$ neither $\mathcal{P}_{\mathrm{c}}$.

Lemma 2.2.1. *Let $\mathfrak{G}$ be an alphabet, $\mathcal{P}$ be a subset of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$, $\mathrm{a} \in \mathfrak{G}$, and $\mathcal{S}$ be a $\mathcal{P}_{\mathrm{a}}$-consistent word. If $\mathfrak{t}$ is an $\mathcal{S}$-admissible $\mathfrak{G}$-tree, then $\mathfrak{t}$ prefix-avoids $\mathcal{P}_{\mathrm{a}}$.*

Proof. Let us denote by k the arity of a. Since $\mathfrak{t}$ is $\mathcal{S}$-admissible, for all $i \in [k]$ and $\mathfrak{s} \in \mathcal{S}_i$, we have $\mathfrak{s} \not\preccurlyeq_{\mathrm{p}} \mathfrak{t}(i)$. Since for any $\mathfrak{r} \in \mathcal{P}_{\mathrm{a}}$, there is a $j \in [k]$ such that $\mathfrak{r}(j) \neq |$ and $\mathfrak{r}(j) \in \mathcal{S}_j$, we have in particular that $\mathfrak{r}(j) \not\preccurlyeq_{\mathrm{p}} \mathfrak{t}(j)$. Since moreover the root of $\mathfrak{t}$ is labeled by a, by Lemma 2.1.2, one deduces that $\mathfrak{s} \not\preccurlyeq_{\mathrm{p}} \mathfrak{t}$. □

If $(\mathcal{S}_1, \dots, \mathcal{S}_k)$ and $(\mathcal{S}'_1, \dots, \mathcal{S}'_k)$ are two words of a same length k where each $\mathcal{S}_i$ and $\mathcal{S}'_i$ is a subset of $\mathbf{S}(\mathfrak{G})$, their *sum* is the word

$$(\mathcal{S}_1, \dots, \mathcal{S}_k) \dotplus (\mathcal{S}'_1, \dots, \mathcal{S}'_k) := (\mathcal{S}_1 \cup \mathcal{S}'_1, \dots, \mathcal{S}_k \cup \mathcal{S}'_k). \tag{2.2.5}$$

A $\mathcal{P}_{\mathrm{a}}$-consistent word $(\mathcal{S}_1, \dots, \mathcal{S}_k)$ is *minimal* if any decomposition

$$(\mathcal{S}_1, \dots, \mathcal{S}_k) = (\mathcal{S}'_1, \dots, \mathcal{S}'_k) \dotplus (\mathcal{S}''_1, \dots, \mathcal{S}''_k) \tag{2.2.6}$$

where $(\mathcal{S}'_1, \dots, \mathcal{S}'_k)$ is a $\mathcal{P}_{\mathrm{a}}$-consistent word and $(\mathcal{S}''_1, \dots, \mathcal{S}''_k)$ is a word where each $\mathcal{S}''_i$, $i \in [k]$, is a subset of $\mathbf{S}(\mathfrak{G})$, implies $(\mathcal{S}_1, \dots, \mathcal{S}_k) = (\mathcal{S}'_1, \dots, \mathcal{S}'_k)$. Intuitively, this says that a $\mathcal{P}_{\mathrm{a}}$-consistent word is minimal if the suppression of any tree in one of its letters leads to a word which is not $\mathcal{P}_{\mathrm{a}}$-consistent. We denote by $\mathfrak{M}(\mathcal{P}_{\mathrm{a}})$ the set of all minimal $\mathcal{P}_{\mathrm{a}}$-consistent words.

For instance, by considering the set (2.2.1) of patterns,

$$\mathfrak{M}(\mathcal{P}_{\mathtt{a}}) = \left\{ \left(\mathtt{c}, \emptyset \right) \right\}, \tag{2.2.7a}$$

$$\mathfrak{M}(\mathcal{P}_{\mathtt{b}}) = \{(\emptyset, \emptyset)\}, \tag{2.2.7b}$$

$$\mathfrak{M}(\mathcal{P}_{\mathtt{c}}) = \left\{ \left(\mathtt{a}, \mathtt{b}, \mathtt{a} \right), \left(\mathtt{a} \; \mathtt{b}, \emptyset, \mathtt{a} \right), \left(\mathtt{a} \; \mathtt{b} \; \mathtt{c}(|,|,\mathtt{c}), \emptyset, \emptyset \right) \right\}. \tag{2.2.7c}$$

For any $\mathfrak{G}$-tree, we denote by $\mathrm{Pref}(\mathfrak{t})$ the set of all prefixes of $\mathfrak{t}$.

Lemma 2.2.2. *Let $\mathfrak{G}$ be an alphabet and $\mathcal{P}$ be a subset of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$. If $\mathfrak{t}$ is a $\mathfrak{G}$-tree having its root labeled by $\mathtt{a} \in \mathfrak{G}$ and prefix-avoiding $\mathcal{P}_{\mathtt{a}}$, then there is a minimal $\mathcal{P}_{\mathtt{a}}$-consistent word $\mathcal{S}$ such that $\mathfrak{t}$ is $\mathcal{S}$-admissible.*

Proof. Let us denote by k the arity of $\mathtt{a}$ and let $\mathcal{S} := (\mathcal{S}_1, \dots, \mathcal{S}_k)$ be the word of subsets of $\mathbf{S}(\mathfrak{G})$ defined by $\mathcal{S}_i := \mathbf{S}(\mathfrak{G}) \setminus \mathrm{Pref}(\mathfrak{t}(i))$. Since $\mathfrak{t}$ prefix-avoids $\mathcal{P}_{\mathtt{a}}$, by Lemma 2.1.2, for any $\mathfrak{r} \in \mathcal{P}_{\mathtt{a}}$, there is an $i \in [k]$ such that $\mathfrak{r}(i) \neq |$ and $\mathfrak{r}(i) \not\preccurlyeq_{\mathrm{p}} \mathfrak{t}(i)$. This leads to the fact that $\mathfrak{r}(i) \notin \mathrm{Pref}(\mathfrak{t}(i))$, so that $\mathfrak{r}(i) \in \mathcal{S}_i$. For this reason, $\mathcal{S}$ is $\mathcal{P}_{\mathtt{a}}$-consistent. Moreover, it follows directly from the definition of $\mathcal{S}$ that $\mathfrak{t}$ is $\mathcal{S}$-admissible. Finally, by definition of minimal $\mathcal{P}_{\mathtt{a}}$-consistent words, there exists a minimal $\mathcal{P}_{\mathtt{a}}$-consistent word $\mathcal{S}' := (\mathcal{S}'_1, \dots, \mathcal{S}'_k)$ such that $\mathcal{S}'_i \subseteq \mathcal{S}_i$ for all $i \in [k]$. The statement of the lemma follows. □

By combining Lemmas 2.2.1 and 2.2.2 together, it follows that for any subset $\mathcal{P}$ of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$ and any letter $\mathtt{a} \in \mathfrak{G}$, a $\mathfrak{G}$-tree $\mathfrak{t}$ having its root labeled by $\mathtt{a}$ prefix-avoids $\mathcal{P}$ if and only if there exists a minimal $\mathcal{P}_{\mathtt{a}}$-consistent word such that $\mathfrak{t}$ is $\mathcal{S}$-admissible.

Lemma 2.2.3. *Let $\mathfrak{G}$ be an alphabet, $\mathcal{P}$ and $\mathcal{Q}$ be two subsets of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$, and $\mathfrak{t}$ be a $\mathfrak{G}$-tree having its root labeled by $\mathtt{a} \in \mathfrak{G}(k)$. Then, $\mathfrak{t}$ factor-avoids $\mathcal{P}$ and prefix-avoids $\mathcal{Q}$ if and only if for all $i \in [k]$, $\mathfrak{t}(i)$ factor-avoid $\mathcal{P}$ and there exists a minimal $(\mathcal{P} \cup \mathcal{Q})_{\mathtt{a}}$-consistent word $\mathcal{S}$ such that $\mathfrak{t}$ is $\mathcal{S}$-admissible.*

Proof. Assume that $\mathfrak{t}$ factor-avoids $\mathcal{P}$ and prefix-avoids $\mathcal{Q}$. The fact that $\mathfrak{t}$ factor-avoids $\mathcal{P}$ implies in particular that $\mathfrak{t}$ prefix-avoids $\mathcal{P}$ (see Proposition 2.1.1). Hence, $\mathfrak{t}$ prefix-avoids $\mathcal{P} \cup \mathcal{Q}$. Now, by Lemma 2.2.2, and since the root of $\mathfrak{t}$ is labeled by $\mathtt{a}$, there exists a minimal $(\mathcal{P} \cup \mathcal{Q})_{\mathtt{a}}$-consistent word $\mathcal{S}$ such that $\mathfrak{t}$ is $\mathcal{S}$-admissible. Conversely, assume that for all $i \in [k]$, $\mathfrak{t}(i)$ factor-avoid $\mathcal{P}$ and that there exists a minimal $(\mathcal{P} \cup \mathcal{Q})_{\mathtt{a}}$-consistent word $\mathcal{S}$ such that $\mathfrak{t}$ is $\mathcal{S}$-admissible. By Lemma 2.2.1, $\mathfrak{t}$ prefix-avoids $(\mathcal{P} \cup \mathcal{Q})_{\mathtt{a}}$. Therefore, since $\mathfrak{t}$ prefix-avoids $\mathcal{P}$ and since for each $i \in [k]$, $\mathfrak{t}(i)$ factor-avoids $\mathcal{P}$, we have that $\mathfrak{t}$ factor-avoids $\mathcal{P}$. Since moreover $\mathfrak{t}$ prefix-avoids $\mathcal{Q}$, we finally have that $\mathfrak{t}$ factor-avoids $\mathcal{P}$ and prefix-avoids $\mathcal{Q}$. □

2.2.2. *Equations for tree series.* For any subsets $\mathcal{P}$ and $\mathbb{Q}$ of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$, let $\mathbb{P}_{\mathcal{P},\mathbb{Q}}$ be the predicate on $\mathbf{S}(\mathfrak{G})$ wherein $\mathbb{P}_{\mathcal{P},\mathbb{Q}}(\mathrm{t})$ holds if and only if $\mathcal{P} \not\preccurlyeq_{\mathrm{f}} \mathrm{t}$ and $\mathbb{Q} \not\preccurlyeq_{\mathrm{p}} \mathrm{t}$. Let also $\mathbf{F}(\mathcal{P}, \mathbb{Q})$ be the $\mathfrak{G}$-tree series defined by

$$\mathbf{F}(\mathcal{P}, \mathbb{Q}) := \mathrm{pr}\left(\mathbb{P}_{\mathcal{P},\mathbb{Q}}\right). \tag{2.2.8}$$

In other terms, $\mathbf{F}(\mathcal{P}, \mathbb{Q})$ is the characteristic series of all $\mathfrak{G}$-trees factor-avoiding all trees of $\mathcal{P}$ and prefix-avoiding all trees of $\mathbb{Q}$. Since $\mathbf{F}(\mathcal{P}, \emptyset) = \mathbf{F}(\mathcal{P})$, we can regard $\mathbf{F}(\mathcal{P}, \mathbb{Q})$ as a refinement of $\mathbf{F}(\mathcal{P})$. Observe also that $\mathbf{F}(\mathcal{P}, \mathcal{P}') = \mathbf{F}(\mathcal{P})$ for all subsets $\mathcal{P}'$ of $\mathcal{P}$. As a side remark, observe that $\mathbf{F}(\emptyset, \mathbb{Q})$ is the characteristic series of the $\mathfrak{G}$-trees prefix-avoiding $\mathbb{Q}$.

Theorem 2.2.4. *Let $\mathfrak{G}$ be an alphabet, and $\mathcal{P}$ and $\mathbb{Q}$ be two subsets of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$ such that for any $\mathrm{a} \in \mathfrak{G}$, there are finitely many minimal $(\mathcal{P} \cup \mathbb{Q})_{\mathrm{a}}$-consistent words. The $\mathfrak{G}$-tree series $\mathbf{F}(\mathcal{P}, \mathbb{Q})$ satisfies*

$$\mathbf{F}(\mathcal{P}, \mathbb{Q}) = | + \sum_{\substack{k \geqslant 1 \\ \mathrm{a} \in \mathfrak{G}(k)}} \sum_{\substack{\ell \geqslant 1 \\ \left\{\mathcal{R}^{(1)}, \dots, \mathcal{R}^{(\ell)}\right\} \subseteq \mathfrak{M}((\mathcal{P} \cup \mathbb{Q})_{\mathrm{a}}) \\ (\mathcal{S}_1, \dots, \mathcal{S}_k) = \mathcal{R}^{(1)} \dotplus \cdots \dotplus \mathcal{R}^{(\ell)}}} (-1)^{1+\ell} \, \mathrm{a} \,\bar{\circ}\, \left[\mathbf{F}(\mathcal{P}, \mathcal{S}_1), \dots, \mathbf{F}(\mathcal{P}, \mathcal{S}_k)\right]. \tag{2.2.9}$$

Proof. For any $\mathrm{a} \in \mathfrak{G}(k)$ and any $\mathcal{S} \in \mathfrak{M}((\mathcal{P} \cup \mathbb{Q})_{\mathrm{a}})$, let $\mathbb{P}_{\mathrm{a},\mathcal{S}}$ be the predicate on $\mathbf{S}(\mathfrak{G})$ wherein $\mathbb{P}_{\mathrm{a},\mathcal{S}}(\mathrm{t})$ holds if and only if $\mathcal{P} \not\preccurlyeq_{\mathrm{f}} \mathrm{t}$, $\mathbb{Q} \not\preccurlyeq_{\mathrm{p}} \mathrm{t}$, and t is $\mathcal{S}$-admissible. As a consequence of Lemma 2.2.3, we have

$$\mathrm{pr}\left(\mathbb{P}_{\mathrm{a},\mathcal{S}}\right) = \mathrm{a} \,\bar{\circ}\, \left[\mathbf{F}(\mathcal{P}, \mathcal{S}_1), \dots, \mathbf{F}(\mathcal{P}, \mathcal{S}_k)\right]. \tag{2.2.10}$$

Now, observe that for any $\mathcal{S}, \mathcal{S}' \in \mathfrak{M}((\mathcal{P} \cup \mathbb{Q})_{\mathrm{a}})$, the predicates $\mathbb{P}_{\mathrm{a},\mathcal{S} \dotplus \mathcal{S}'}$ and $\mathbb{P}_{\mathrm{a},\mathcal{S}} \wedge \mathbb{P}_{\mathrm{a},\mathcal{S}'}$ are equal. Observe also that the characteristic series $\mathbf{f}_{\mathrm{a}}$ of the $\mathfrak{G}$-trees factor-avoiding $\mathcal{P}$, prefix-avoiding $\mathbb{Q}$, and with a root labeled by a, satisfies

$$\mathbf{f}_{\mathrm{a}} = \mathrm{pr}\left(\bigvee_{\mathcal{S} \in \mathfrak{M}((\mathcal{P} \cup \mathbb{Q})_{\mathrm{a}})} \mathbb{P}_{\mathrm{a},\mathcal{S}}\right). \tag{2.2.11}$$

Since, by hypothesis, $\mathfrak{M}((\mathcal{P} \cup \mathbb{Q})_{\mathrm{a}})$ is finite, these three previous properties lead, by using Lemma 1.2.1, to the relation

$$\mathbf{f}_{\mathrm{a}} = \sum_{\substack{\ell \geqslant 1 \\ \left\{\mathcal{R}^{(1)}, \dots, \mathcal{R}^{(\ell)}\right\} \subseteq \mathfrak{M}((\mathcal{P} \cup \mathbb{Q})_{\mathrm{a}}) \\ (\mathcal{S}_1, \dots, \mathcal{S}_k) = \mathcal{R}^{(1)} \dotplus \cdots \dotplus \mathcal{R}^{(\ell)}}} (-1)^{1+\ell} \, \mathrm{a} \,\bar{\circ}\, \left[\mathbf{F}(\mathcal{P}, \mathcal{S}_1), \dots, \mathbf{F}(\mathcal{P}, \mathcal{S}_k)\right]. \tag{2.2.12}$$

Finally, since any tree factor-avoiding $\mathcal{P}$ and prefix-avoiding $\mathbb{Q}$ can be either empty of have a root labeled by a for any $\mathrm{a} \in \mathfrak{G}$, we have

$$\mathbf{F}(\mathcal{P}, \mathbb{Q}) = | + \sum_{\mathrm{a} \in \mathfrak{G}} \mathbf{f}_{\mathrm{a}}. \tag{2.2.13}$$

This last relation shows that (2.2.9) holds. □

Let us consider an example of expression brought by Theorem 2.2.4 by considering the set (2.2.1) of patterns. We have

$$\mathbf{F}(\mathcal{P}, \emptyset) = | + \mathrm{a} \,\bar{\circ}\, \left[\mathbf{F}\left(\mathcal{P}, \begin{smallmatrix} | \\ \mathrm{c} \\ /|\backslash \end{smallmatrix}\right), \mathbf{F}(\mathcal{P}, \emptyset)\right] + \mathrm{b} \,\bar{\circ}\, \left[\mathbf{F}(\mathcal{P}, \emptyset), \mathbf{F}(\mathcal{P}, \emptyset)\right]$$

$$
\begin{aligned}
&+ \mathrm{c}\,\bar{\mathrm{o}} \left[\mathbf{F}\left(\mathcal{P}, \mathrm{a}\right), \mathbf{F}\left(\mathcal{P}, \mathrm{b}\right), \mathbf{F}\left(\mathcal{P}, \mathrm{a}\right)\right] + \mathrm{c}\,\bar{\mathrm{o}} \left[\mathbf{F}\left(\mathcal{P}, \mathrm{a}\ \mathrm{b}\right), \mathbf{F}\left(\mathcal{P}, \emptyset\right), \mathbf{F}\left(\mathcal{P}, \mathrm{a}\right)\right] \\
&+ \mathrm{c}\,\bar{\mathrm{o}} \left[\mathbf{F}\left(\mathcal{P}, \mathrm{a}\ \mathrm{b}\ \mathrm{c}(|,|,\mathrm{c})\right), \mathbf{F}\left(\mathcal{P}, \emptyset\right), \mathbf{F}\left(\mathcal{P}, \emptyset\right)\right] - \mathrm{c}\,\bar{\mathrm{o}} \left[\mathbf{F}\left(\mathcal{P}, \mathrm{a}\ \mathrm{b}\right), \mathbf{F}\left(\mathcal{P}, \mathrm{b}\right), \mathbf{F}\left(\mathcal{P}, \mathrm{a}\right)\right] \\
&- \mathrm{c}\,\bar{\mathrm{o}} \left[\mathbf{F}\left(\mathcal{P}, \mathrm{a}\ \mathrm{b}\ \mathrm{c}(|,|,\mathrm{c})\right), \mathbf{F}\left(\mathcal{P}, \mathrm{b}\right), \mathbf{F}\left(\mathcal{P}, \mathrm{a}\right)\right] - \mathrm{c}\,\bar{\mathrm{o}} \left[\mathbf{F}\left(\mathcal{P}, \mathrm{a}\ \mathrm{b}\ \mathrm{c}(|,|,\mathrm{c})\right), \mathbf{F}\left(\mathcal{P}, \emptyset\right), \mathbf{F}\left(\mathcal{P}, \mathrm{a}\right)\right] \\
&+ \mathrm{c}\,\bar{\mathrm{o}} \left[\mathbf{F}\left(\mathcal{P}, \mathrm{a}\ \mathrm{b}\ \mathrm{c}(|,|,\mathrm{c})\right), \mathbf{F}\left(\mathcal{P}, \mathrm{b}\right), \mathbf{F}\left(\mathcal{P}, \mathrm{a}\right)\right]. \qquad (2.2.14)
\end{aligned}
$$

Observe that the last term of (2.2.14) is the opposite of the antepenultimate term so that they annihilate.

2.3. **Properties and applications.** Consequences of Theorem 2.2.4 are now presented. In particular, we explain how to obtain a system of equations of generating series to enumerate the syntax trees factor-avoiding a set $\mathcal{P}$ of patterns and prefix-avoiding a set $\mathcal{Q}$ of patterns. We also apply the aforementioned result for particular sets of patterns consisting in stringy trees.

2.3.1. *Systems of equations.* Given two subsets $\mathcal{P}$ and $\mathcal{Q}$ of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$ satisfying the conditions of Theorem 2.2.4, one can express the series $\mathbf{F}(\mathcal{P}, \mathcal{Q})$ through (2.2.9). Some other series $\mathbf{F}(\mathcal{P}, \mathcal{S}_i)$ could appear in the expression, and these series can themselves be expressed through (2.2.9) when the conditions of the theorem are satisfied. When it is the case, Theorem 2.2.4 leads to a (possibly infinite) system of equations describing the series $\mathbf{F}(\mathcal{P}, \mathcal{Q})$, called *system* of $\mathbf{F}(\mathcal{P}, \mathcal{Q})$.

Lemma 2.3.1. *Let $\mathfrak{G}$ be an alphabet, $\mathcal{P}$ be a subset of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$, and $\mathrm{a} \in \mathfrak{G}(k)$. If $\mathcal{P}_{\mathrm{a}}$ is finite, then the set of all minimal $\mathcal{P}_{\mathrm{a}}$-consistent words is finite and its cardinality is no greater than $k^{\#\mathcal{P}_{\mathrm{a}}}$.*

Proof. We proceed by induction on the cardinality ℓ of $\mathcal{P}_{\mathrm{a}}$. If $\ell = 0$, the only $\mathcal{P}_{\mathrm{a}}$-consistent word is the word $(\mathcal{S}_1, \dots, \mathcal{S}_k)$ such that $\mathcal{S}_i := \emptyset$ for all $i \in [k]$. Hence, the statement of the lemma holds in this case. Assume now that the statement of the lemma holds when $\mathcal{P}_{\mathrm{a}}$ has cardinality ℓ. Let $\mathfrak{s}$ be a $\mathfrak{G}$-tree having its root labeled by a. If $\mathcal{S} := (\mathcal{S}_1, \dots, \mathcal{S}_k)$ is a $\mathcal{P}_{\mathrm{a}}$-consistent word, when $j \in [k]$ is an index such that $\mathfrak{s}(j) \neq |$, let us denote by $\mathcal{S}^{(j)} := (\mathcal{S}'_1, \dots, \mathcal{S}'_k)$ the word defined by $\mathcal{S}'_j := \mathcal{S}_j \cup \{\mathfrak{s}(j)\}$ and $\mathcal{S}'_i := \mathcal{S}_i$ for any $i \in [k] \setminus \{j\}$. By construction, $\mathcal{S}^{(j)}$ is a minimal $(\mathcal{P} \cup \{\mathfrak{s}\})_{\mathrm{a}}$-consistent word and there are at most k such words. By induction hypothesis, there are at most k^ℓ minimal $\mathcal{P}_{\mathrm{a}}$-consistent words and therefore, at most $k^{\ell+1}$ minimal $(\mathcal{P} \cup \{\mathfrak{s}\})_{\mathrm{a}}$-consistent words. □

For any $\mathfrak{G}$-tree, we denote by $\mathrm{Suff}(\mathfrak{t})$ the set of all suffixes of $\mathfrak{t}$.

Proposition 2.3.2. *Let $\mathfrak{G}$ be an alphabet, and $\mathcal{P}$ and $\mathcal{Q}$ be two subsets of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$. If $\mathcal{P}$ and $\mathcal{Q}$ are finite, then the system of $\mathbf{F}(\mathcal{P}, \mathcal{Q})$ is well-defined and contains finitely many equations.*

Proof. Let $\mathsf{a} \in \mathfrak{G}(k)$. Since $\mathcal{P}$ and $\mathcal{Q}$ are finite, $(\mathcal{P} \cup \mathcal{Q})_{\mathsf{a}}$ is finite. Therefore, by Lemma 2.3.1, $\mathfrak{M}\left((\mathcal{P} \cup \mathcal{Q})_{\mathsf{a}}\right)$ is finite. Moreover, any minimal $(\mathcal{P} \cup \mathcal{Q})_{\mathsf{a}}$-consistent word $(\mathcal{S}_1, \dots, \mathcal{S}_k)$ is such that each $\mathcal{S}_i$, $i \in [k]$, contains only suffixes of trees of $(\mathcal{P} \cup \mathcal{Q})_{\mathsf{a}}$. For this reason, all terms $\mathbf{F}(\mathcal{P}, \mathcal{S}_i)$ appearing in the equation (2.2.9) of $\mathbf{F}(\mathcal{P}, \mathcal{Q})$ satisfy

$$\mathcal{S}_i \subseteq \bigcup_{\mathsf{t} \in \mathcal{P} \cup \mathcal{Q}} \mathrm{Suff}(\mathsf{t}). \tag{2.3.1}$$

Since any $\mathfrak{G}$-tree has a finite number of suffixes, there are finitely many sets $\mathcal{S}_i$ satisfying (2.3.1). The statement of the proposition follows. □

2.3.2. *Limits.* Let $\mathcal{P}$ be a subset of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$. For any integer $d \geqslant 0$, let

$$\mathcal{P}_{|d} := \{\mathsf{t} \in \mathcal{P} : \deg(\mathsf{t}) \leqslant d\}. \tag{2.3.2}$$

In other words, $\mathcal{P}_{|d}$ is the subset of $\mathcal{P}$ of the patterns having degrees no greater than d.

Proposition 2.3.3. *Let $\mathfrak{G}$ be an alphabet, and $\mathcal{P}$ and $\mathcal{Q}$ be two subsets of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$. Then,*

$$\lim_{d \to \infty} \mathbf{F}\left(\mathcal{P}_{|d}, \mathcal{Q}_{|d}\right) = \mathbf{F}(\mathcal{P}, \mathcal{Q}). \tag{2.3.3}$$

Proof. Since any $\mathfrak{G}$-tree t factor-avoids (resp. prefix-avoids) all patterns of degrees greater than $\deg(\mathsf{t})$, for any $d \geqslant \deg(\mathsf{t})$,

$$\langle \mathsf{t}, \mathbf{F}(\mathcal{P}, \mathcal{Q}) \rangle = \left\langle \mathsf{t}, \mathbf{F}\left(\mathcal{P}_{|d}, \mathcal{Q}_{|d}\right) \right\rangle. \tag{2.3.4}$$

This implies that the coefficients of the series $\mathbf{F}(\mathcal{P}, \mathcal{Q})$ and $\mathbf{F}\left(\mathcal{P}_{|d}, \mathcal{Q}_{|d}\right)$ coincide for all the $\mathfrak{G}$-trees of degrees no greater than d. The statement of the proposition follows. □

Theorem 2.2.4 and Proposition 2.3.3 allow us together to obtain systems of equations for $\mathbf{F}(\mathcal{P}, \mathcal{Q})$ even when $\mathcal{P}$ and $\mathcal{Q}$ are infinite subsets of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$ that do not satisfy the hypothesis of Theorem 2.2.4.

2.3.3. *Generating series and systems of equations.* For any subset $\mathcal{P}$ of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$, let $F(\mathcal{P})$ be the series of $\mathbb{K}\langle\langle t, q, \mathbf{Q}_{\mathfrak{G}} \rangle\rangle$ defined by $F(\mathcal{P}) := \mathrm{en}(\mathbf{F}(\mathcal{P}))$. In the same way, for any subsets $\mathcal{P}$ and $\mathcal{Q}$ of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$, let $F(\mathcal{P}, \mathcal{Q})$ be the series of $\mathbb{K}\langle\langle t, q, \mathbf{Q}_{\mathfrak{G}} \rangle\rangle$ defined by $F(\mathcal{P}, \mathcal{Q}) := \mathrm{en}(\mathbf{F}(\mathcal{P}, \mathcal{Q}))$. The series $F(\mathcal{P})$ is the generating series of the set of the $\mathfrak{G}$-trees factor-avoiding $\mathcal{P}$, and $F(\mathcal{P}, \mathcal{Q})$ is the generating series of the set of the $\mathfrak{G}$-trees factor-avoiding $\mathcal{P}$ and prefix-avoiding $\mathcal{Q}$.

Proposition 2.3.4. *Let $\mathfrak{G}$ be an alphabet, and $\mathcal{P}$ and $\mathcal{Q}$ be two subsets of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$ such that for any $\mathsf{a} \in \mathfrak{G}$, $(\mathcal{P} \cup \mathcal{Q})_{\mathsf{a}}$ is finite. The generating series $F(\mathcal{P}, \mathcal{Q})$ satisfies*

$$F(\mathcal{P}, \mathcal{Q}) = t + q \sum_{\substack{k \geqslant 1 \\ \mathsf{a} \in \mathfrak{G}(k)}} q_{\mathsf{a}} \sum_{\substack{\ell \geqslant 1 \\ \left\{\mathcal{R}^{(1)}, \dots, \mathcal{R}^{(\ell)}\right\} \subseteq \mathfrak{M}((\mathcal{P} \cup \mathcal{Q})_{\mathsf{a}}) \\ (\mathcal{S}_1, \dots, \mathcal{S}_k) = \mathcal{R}^{(1)} \dotplus \cdots \dotplus \mathcal{R}^{(\ell)}}} (-1)^{1+\ell} \prod_{i \in [k]} F(\mathcal{P}, \mathcal{S}_i). \tag{2.3.5}$$

Proof. Relation (2.3.5) is obtained by considering the enumerative images of the left and right members of (2.2.9) provided by Theorem 2.2.4, together with Proposition 1.2.2. □

2.3.4. *Avoiding stringy trees.* A $\mathfrak{G}$-tree $\mathfrak{t}$ is *stringy* if the height of $\mathfrak{t}$ is equal to the degree of $\mathfrak{t}$. This is equivalent to the fact that any internal node of $\mathfrak{t}$ has at most one child being an internal node.

For any set $\mathcal{P}$ of $\mathfrak{G}$-trees, $\mathsf{a} \in \mathfrak{G}(k)$, and $i \in [k]$, let

$$\partial_{\mathsf{a},i}(\mathcal{P}) := \{\mathfrak{s} \in \mathbf{S}(\mathfrak{G}) : \mathsf{a} \circ_i \mathfrak{s} \in \mathcal{P}\}. \tag{2.3.6}$$

In other words, $\partial_{\mathsf{a},i}(\mathcal{P})$ is the set of the $\mathfrak{G}$-trees obtained by keeping the ith subtrees of the trees whose roots are labeled by a in $\mathcal{P}$.

Proposition 2.3.5. *Let $\mathfrak{G}$ be an alphabet and $\mathcal{P}$ and $\mathcal{Q}$ be two subsets of $\mathbf{S}(\mathfrak{G}) \setminus \{|\}$ consisting only in stringy trees. The $\mathfrak{G}$-tree series $\mathbf{F}(\mathcal{P}, \mathcal{Q})$ satisfies*

$$\mathbf{F}(\mathcal{P}, \mathcal{Q}) = | + \sum_{\substack{k \geqslant 1 \\ \mathsf{a} \in \mathfrak{G}(k) \\ \mathfrak{c}(\mathsf{a}) \notin \mathcal{P} \cup \mathcal{Q}}} \mathsf{a} \,\bar{\circ} \left[\mathbf{F}\left(\mathcal{P}, \partial_{\mathsf{a},1}(\mathcal{P} \cup \mathcal{Q})\right), \dots, \mathbf{F}\left(\mathcal{P}, \partial_{\mathsf{a},k}(\mathcal{P} \cup \mathcal{Q})\right)\right]. \tag{2.3.7}$$

Proof. Let $\mathsf{a} \in \mathfrak{G}(k)$. When $\mathfrak{c}(\mathsf{a})$ is in $\mathcal{P} \cup \mathcal{Q}$, by definition of consistent words, there is no $(\mathcal{P} \cup \mathcal{Q})_{\mathsf{a}}$-consistent word. When $\mathfrak{c}(\mathsf{a})$ is not in $\mathcal{P} \cup \mathcal{Q}$, by definition of minimal consistent words, the only minimal $(\mathcal{P} \cup \mathcal{Q})_{\mathsf{a}}$-consistent word is the word $\mathcal{S} := (\mathcal{S}_1, \dots, \mathcal{S}_k)$ where $\mathcal{S}_i := \partial_{\mathsf{a},i}(\mathcal{P} \cup \mathcal{Q})$ for any $i \in [k]$. Now, (2.3.7) is a consequence of Theorem 2.2.4. □

Let us call *$\mathfrak{G}$-word* any $\mathfrak{G}$-tree where $\mathfrak{G}$ is an alphabet concentrated in arity 1. This designation is justified by the fact that one can encode any word $\mathsf{a}_1 \dots \mathsf{a}_d$ on $\mathfrak{G}$ through the tree $\mathsf{a}_1 \circ_1 \cdots \circ_1 \mathsf{a}_d$. When $\mathcal{P}$ contains only $\mathfrak{G}$-words different from the leaf, $\mathcal{P}$ specifies forbidden configurations of word factors. Since a $\mathfrak{G}$-word is obviously stringy, Proposition 2.3.5 provides in this context a system of equations to describe the series of words avoiding factors. This problem consisting in enumerating words avoiding as factors a given set was originally stated and solved in [GJ79] (see also [NZ99]).

Besides, when $\mathfrak{G}$ is any alphabet, let us call *$\mathfrak{G}$-edge* any $\mathfrak{G}$-tree of degree 2. This appellation is justified by the fact that any tree of degree 2 contains exactly one edge connecting two internal nodes. When $\mathcal{P}$ contains only $\mathfrak{G}$-edges, $\mathcal{P}$ specifies forbidden configurations of edges. Since a $\mathfrak{G}$-edge is obviously stringy, Proposition 2.3.5 provides in this context a system of equations to describe the series of trees avoiding edges. This particular case of pattern avoidance in trees was studied in [Lod05] (see also [Par93]).

2.3.5. *Sets of patterns for some algebraic series.* Let us assume here that $\mathbb{K}$ is the field $\mathbb{Q}$. A series $\mathbf{f}$ of $\mathbb{K}\langle\langle t \rangle\rangle$ is *$\mathbb{N}$-algebraic* if $\mathbf{f}$ satisfies the equation

$$\mathbf{f} = \sum_{0 \leqslant n \leqslant d} P_n \, \mathbf{f}^n \tag{2.3.8}$$

where d is a certain nonnegative integer, for all $0 \leqslant n \leqslant d$, the P_n are polynomials of $\mathbb{Q}\langle t \rangle$ having all coefficients in $\mathbb{N}$, and $\left\langle t^0, P_1 \right\rangle = 0$. For instance, the series $\mathbf{f}$ satisfying

$$\mathbf{f} = t + t^3 + \left(t + t^2\right)\mathbf{f} + \left(1 + 2t^3\right)\mathbf{f}^2 \tag{2.3.9}$$

is $\mathbb{N}$-algebraic.

Proposition 2.3.6. *Let* $\mathbf{f}$ *be an* $\mathbb{N}$*-algebraic series of the form* (2.3.8) *such that* $\langle t^0, P_0 \rangle = 0$ *and* $\langle t^1, P_0 \rangle = 1$*. Let the alphabet* $\mathfrak{G} := \bigsqcup_{n \geqslant 2} \mathfrak{G}(n)$ *where, for any* $n \geqslant 2$*,*

$$\mathfrak{G}(n) := \bigsqcup_{\substack{k,\ell \geqslant 0 \\ k+\ell = n}} \left\{ \mathrm{a}_{k,\ell}^{(m)} : 1 \leqslant m \leqslant \langle t^\ell, P_k \rangle \right\} \tag{2.3.10}$$

and the set of patterns

$$\mathcal{P} := \bigsqcup_{\substack{\mathrm{a}_{k,\ell}^{(m)} \in \mathfrak{G} \\ i \in [\ell]}} \left\{ \mathrm{a}_{k,\ell}^{(m)} \circ_i \mathrm{b} : \mathrm{b} \in \mathfrak{G} \right\}. \tag{2.3.11}$$

The specialization $F(\mathcal{P}, \emptyset)_{|q:=1, q_{\mathrm{a}}:=1, \mathrm{a} \in \mathfrak{G}}$ *satisfies the same algebraic equation as the one satisfied by* $\mathbf{f}$*.*

Proof. Observe that $\mathcal{P}$ contains only stringy trees. Therefore, the characteristic series $\mathbf{F}(\mathcal{P}, \emptyset)$ of the trees factor-avoiding $\mathcal{P}$ is described by Proposition 2.3.5 and satisfies

$$\mathbf{F}(\mathcal{P}, \emptyset) = | + \sum_{\mathrm{a}_{k,\ell}^{(m)} \in \mathfrak{G}} \mathrm{a}_{k,\ell}^{(m)} \,\bar{\circ}\, \Big[\underbrace{\mathbf{F}(\mathcal{P}, \mathcal{Q}), \dots, \mathbf{F}(\mathcal{P}, \mathcal{Q})}_{\times \ell}, \underbrace{\mathbf{F}(\mathcal{P}, \emptyset), \dots, \mathbf{F}(\mathcal{P}, \emptyset)}_{\times k} \Big] \tag{2.3.12}$$

where $\mathcal{Q} := \{\mathfrak{c}(\mathrm{b}) : \mathrm{b} \in \mathfrak{G}\}$. Now, due to Proposition 2.3.4, the enumerative image $F(\mathcal{P}, \emptyset)$ of $\mathbf{F}(\mathcal{P}, \emptyset)$ satisfies

$$F(\mathcal{P}, \emptyset) = t + q \sum_{\mathrm{a}_{k,\ell}^{(m)} \in \mathfrak{G}} q_{\mathrm{a}_{k,\ell}^{(m)}} F(\mathcal{P}, \mathcal{Q})^\ell \, F(\mathcal{P}, \emptyset)^k \tag{2.3.13}$$

where $F(\mathcal{P}, \mathcal{Q}) = t$. The statement of the proposition follows. □

Observe that the alphabet $\mathfrak{G}$ provided by Proposition 2.3.6 has

$$\#\mathfrak{G} = \sum_{\substack{k,\ell \geqslant 0 \\ k+\ell \geqslant 2}} \langle t^\ell, P_k \rangle \tag{2.3.14}$$

letters, and the set $\mathcal{P}$ is made of

$$\#\mathcal{P} = (\#\mathfrak{G}) \sum_{\substack{k \geqslant 0, \ell \geqslant 0 \\ k+\ell \geqslant 2}} \langle t^\ell, P_k \rangle \, \ell \tag{2.3.15}$$

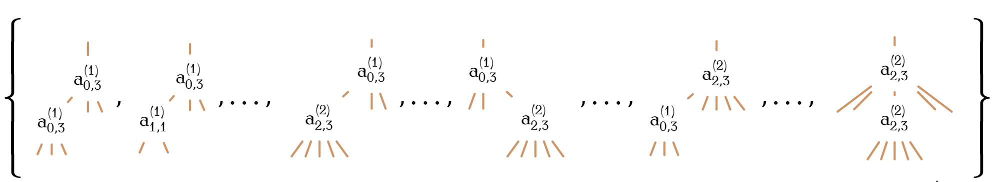
patterns.

Let us consider for example the series $\mathbf{f}$ of (2.3.9). The alphabet and set of patterns specified by Proposition 2.3.6, are

$$\mathfrak{G} := \left\{ \mathrm{a}_{0,3}^{(1)}, \mathrm{a}_{1,1}^{(1)}, \mathrm{a}_{1,2}^{(1)}, \mathrm{a}_{2,0}^{(1)}, \mathrm{a}_{2,3}^{(1)}, \mathrm{a}_{2,3}^{(2)} \right\} \tag{2.3.16}$$

and

$\mathcal{P} :=$



(2.3.17)

The cardinal of $\mathcal{P}$ is $6 \times (1 \times 3 + 1 \times 1 + 1 \times 2 + 2 \times 3) = 72$.

2.3.6. *Examples.* Let us consider some complete examples of systems.

• *Example 1.* Let the alphabet $\mathfrak{G} := \mathfrak{G}(2) := \{\mathsf{a}_i : i \in \mathbb{N}\}$ and the set of patterns

$$\mathcal{P} := \left\{ \begin{smallmatrix} & & \mathsf{a}_i & \\ & \mathsf{a}_i & & \\ / & & \backslash & \end{smallmatrix} : i \in \mathbb{N} \right\}. \tag{2.3.18}$$

By Proposition 2.3.5, we obtain the system

$$\mathbf{F}(\mathcal{P}, \emptyset) = | + \sum_{i \in \mathbb{N}} \mathsf{a}_i \,\bar{\circ}\, \left[\mathbf{F}\left(\mathcal{P}, \begin{smallmatrix} \mathsf{a}_i \\ / \ \backslash \end{smallmatrix}\right), \mathbf{F}(\mathcal{P}, \emptyset)\right], \tag{2.3.19a}$$

$$\mathbf{F}\left(\mathcal{P}, \begin{smallmatrix} \mathsf{a}_i \\ / \ \backslash \end{smallmatrix}\right) = | + \sum_{\substack{j \in \mathbb{N} \\ j \neq i}} \mathsf{a}_j \,\bar{\circ}\, \left[\mathbf{F}\left(\mathcal{P}, \begin{smallmatrix} \mathsf{a}_j \\ / \ \backslash \end{smallmatrix}\right), \mathbf{F}(\mathcal{P}, \emptyset)\right], \qquad i \in \mathbb{N}, \tag{2.3.19b}$$

for the $\mathfrak{G}$-trees factor-avoiding $\mathcal{P}$. Observe that we work here with an infinite alphabet and an infinite set of stringy patterns. This system contains an infinite number of equations.

• *Example 2.* Let the alphabet $\mathfrak{G} := \mathfrak{G}(1) := \{\mathsf{a}, \mathsf{b}\}$ and the set of patterns

$$\mathcal{P} := \left\{ \mathsf{a} \circ_1 \underbrace{\mathsf{b} \circ_1 \cdots \circ_1 \mathsf{b}}_{\times k} \circ_1 \mathsf{a} : k \in \mathbb{N} \right\}. \tag{2.3.20}$$

By Proposition 2.3.5, we obtain the system

$$\mathbf{F}(\mathcal{P}, \emptyset) = | + \mathsf{a} \,\bar{\circ}\, [\mathbf{F}(\mathcal{P}, \mathcal{Q})] + \mathsf{b} \,\bar{\circ}\, [\mathbf{F}(\mathcal{P}, \emptyset)], \tag{2.3.21a}$$

$$\mathbf{F}(\mathcal{P}, \mathcal{Q}) = | + \mathsf{b} \,\bar{\circ}\, [\mathbf{F}(\mathcal{P}, \mathcal{Q})], \tag{2.3.21b}$$

for the $\mathfrak{G}$-trees factor-avoiding $\mathcal{P}$, where

$$\mathcal{Q} := \partial_{\mathsf{a},1}(\mathcal{P}) = \left\{ \underbrace{\mathsf{b} \circ_1 \cdots \circ_1 \mathsf{b}}_{\times k} \circ_1 \mathsf{a} : k \in \mathbb{N} \right\}. \tag{2.3.22}$$

Observe that even if $\mathcal{P}$ is an infinite set of stringy patterns, this system contains a finite number of equations. By Proposition 2.3.4, we obtain the system

$$F(\mathcal{P}, \emptyset) = t + q q_{\mathsf{a}} F(\mathcal{P}, \mathcal{Q}) + q q_{\mathsf{b}} F(\mathcal{P}, \emptyset), \tag{2.3.23a}$$

$$F(\mathcal{P}, \mathcal{Q}) = t + q q_{\mathsf{b}} F(\mathcal{P}, \mathcal{Q}), \tag{2.3.23b}$$

for the enumerative image of the characteristic series of the $\mathfrak{G}$-trees factor-avoiding $\mathcal{P}$.

• *Example 3.* Let the alphabet $\mathfrak{G} := \mathfrak{G}(2) := \{\mathsf{a}_i : i \in \mathbb{Z}\}$ and the set of patterns

$$\mathcal{P} := \left\{ \begin{smallmatrix} & & \mathsf{a}_i & \\ & \mathsf{a}_j & & \backslash \\ / & & \backslash & \end{smallmatrix} : i, j \in \mathbb{Z}, j \leqslant i \right\} \cup \left\{ \begin{smallmatrix} & \mathsf{a}_i & & \\ / & & \mathsf{a}_j & \\ & / & & \backslash \end{smallmatrix} : i, j \in \mathbb{Z}, j \leqslant i \right\}. \tag{2.3.24}$$

By a direct inspection of $\mathcal{P}$, it appears that there is a one-to-one correspondence between the set of the trees factor-avoiding $\mathcal{P}$ and the set of *increasing binary trees*, that are binary trees where internal nodes are labeled on $\mathbb{Z}$ in such a way that the label of any node is smaller than the ones of its children. By Proposition 2.3.5, we obtain the system

$$\mathbf{F}(\mathcal{P}, \emptyset) = | + \sum_{i \in \mathbb{Z}} \mathtt{a}_i \,\bar{\circ}\, \left[\mathbf{F}\left(\mathcal{P}, \mathbb{Q}^{(i)}\right), \mathbf{F}\left(\mathcal{P}, \mathbb{Q}^{(i)}\right)\right], \tag{2.3.25a}$$

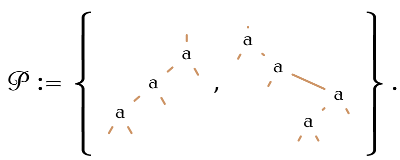
$$\mathbf{F}\left(\mathcal{P}, \mathbb{Q}^{(i)}\right) = | + \sum_{\substack{j \in \mathbb{Z} \\ j \geqslant i+1}} \mathtt{a}_j \,\bar{\circ}\, \left[\mathbf{F}\left(\mathcal{P}, \mathbb{Q}^{(j)}\right), \mathbf{F}\left(\mathcal{P}, \mathbb{Q}^{(j)}\right)\right], \qquad i \in \mathbb{Z}, \tag{2.3.25b}$$

for the $\mathfrak{G}$-trees factor-avoiding $\mathcal{P}$, where for any $j \in \mathbb{Z}$,

$$\mathbb{Q}^{(j)} := \left\{ \mathtt{a}_i \; : i \in \mathbb{Z}, i \leqslant j \right\}. \tag{2.3.26}$$

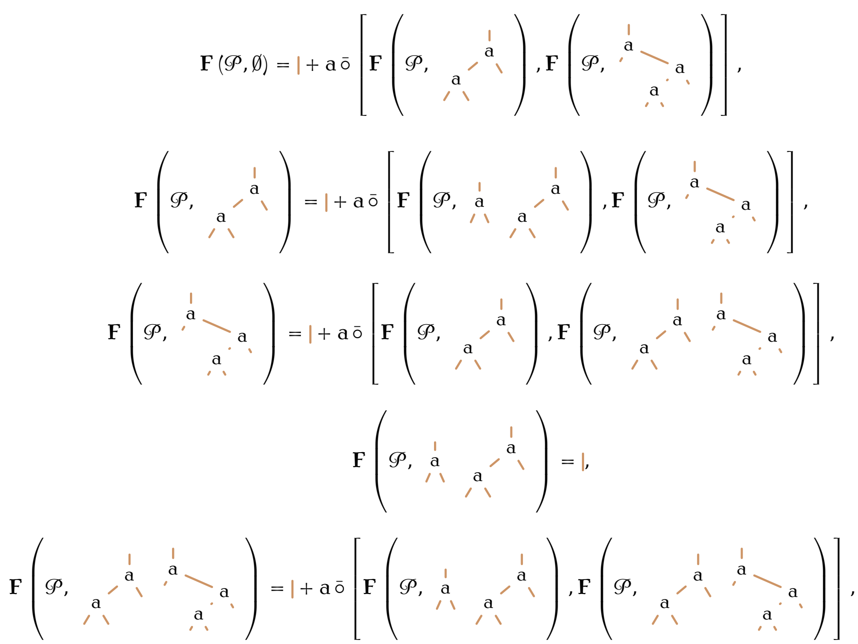
Observe that we work here with an infinite alphabet and an infinite set of stringy patterns. This system contains an infinite number of equations.

• *Example 4.* Let the alphabet $\mathfrak{G} := \mathfrak{G}(2) := \{\mathtt{a}\}$ and the set of patterns



(2.3.27)

By Proposition 2.3.5, we obtain the system



(2.3.28a)
(2.3.28b)
(2.3.28c)
(2.3.28d)
(2.3.28e)

for the $\mathfrak{G}$-trees factor-avoiding $\mathcal{P}$. Observe that we work here with a finite alphabet and a finite set of stringy patterns. The set of patterns considered here comes from an example appearing in [KP15]. Our system shown here is different from the ones presented in this cited work.

• *Example 5.* Let the alphabet $\mathfrak{G} := \mathfrak{G}(2) := \{\mathtt{a}_1, \mathtt{a}_2\}$ and the set of patterns

$$\mathcal{P} := \left\{ \mathtt{a}_2(\mathtt{a}_1, \mathtt{a}_1),\ \mathtt{a}_2(\mathtt{a}_1, \mathtt{a}_2),\ \mathtt{a}_2(\mathtt{a}_2, \mathtt{a}_1),\ \mathtt{a}_2(\mathtt{a}_2, \mathtt{a}_2) \right\}. \tag{2.3.29}$$

A direct inspection of $\mathcal{P}$ shows that a $\mathfrak{G}$-tree factor-avoids $\mathcal{P}$ if and only if any internal node labeled by $\mathtt{a}_2$ have at least a leaf as a child. By Theorem 2.2.4, we obtain the system

$$\begin{aligned} \mathbf{F}(\mathcal{P}, \emptyset) = | + \mathtt{a}_1 \,\bar{\circ}\, [\mathbf{F}(\mathcal{P}, \emptyset), \mathbf{F}(\mathcal{P}, \emptyset)] + \mathtt{a}_2 \,\bar{\circ}\, \left[\mathbf{F}(\mathcal{P}, \emptyset), \mathbf{F}\left(\mathcal{P}, \{\mathtt{a}_1, \mathtt{a}_2\}\right)\right] \\ + \mathtt{a}_2 \,\bar{\circ}\, \left[\mathbf{F}\left(\mathcal{P}, \{\mathtt{a}_1, \mathtt{a}_2\}\right), \mathbf{F}(\mathcal{P}, \emptyset)\right] - \mathtt{a}_2 \,\bar{\circ}\, \left[\mathbf{F}\left(\mathcal{P}, \{\mathtt{a}_1, \mathtt{a}_2\}\right), \mathbf{F}\left(\mathcal{P}, \{\mathtt{a}_1, \mathtt{a}_2\}\right)\right], \end{aligned} \tag{2.3.30a}$$

$$\mathbf{F}\left(\mathcal{P}, \{\mathtt{a}_1, \mathtt{a}_2\}\right) = |, \tag{2.3.30b}$$

for the $\mathfrak{G}$-trees factor-avoiding $\mathcal{P}$. We work here with a finite alphabet and a finite set of non-stringy patterns.

3. Operads, enumeration, and statistics

This section is devoted to use operads as tools to enumerate families of combinatorial objects, jointly with the results presented in he previous sections enumerating trees factor-avoiding some patterns.

3.1. Nonsymmetric set-operads.

We recall here the elementary notions about operads employed thereafter. They mainly come from [Gir18].

3.1.1. *Operad axioms.*

A *nonsymmetric operad in the category of sets*, or a *nonsymmetric operad* for short, is a graded set $\mathcal{O}$ together with maps

$$\circ_i : \mathcal{O}(n) \times \mathcal{O}(m) \to \mathcal{O}(n + m - 1), \qquad 1 \leqslant i \leqslant n, 1 \leqslant m, \tag{3.1.1}$$

called *partial compositions*, and a distinguished element $\mathbb{1} \in \mathcal{O}(1)$, the *unit* of $\mathcal{O}$. This data has to satisfy, for any $x, y, z \in \mathcal{O}$, the three relations

$$(x \circ_i y) \circ_{i+j-1} z = x \circ_i (y \circ_j z), \qquad 1 \leqslant i \leqslant |x|, 1 \leqslant j \leqslant |y|, \tag{3.1.2a}$$

$$(x \circ_i y) \circ_{j+|y|-1} z = (x \circ_j z) \circ_i y, \qquad 1 \leqslant i < j \leqslant |x|, \tag{3.1.2b}$$

$$\mathbb{1} \circ_1 x = x = x \circ_i \mathbb{1}, \qquad 1 \leqslant i \leqslant |x|. \tag{3.1.2c}$$

Since we consider in this work only nonsymmetric operads, we shall call these simply *operads*.

3.1.2. *Elementary definitions.* Given an operad $\mathcal{O}$, one defines the *full composition* maps of $\mathcal{O}$ as the maps

$$\circ : \mathcal{O}(n) \times \mathcal{O}(m_1) \times \cdots \times \mathcal{O}(m_n) \to \mathcal{O}(m_1 + \cdots + m_n), \qquad 1 \leqslant n, 1 \leqslant m_1, \dots, 1 \leqslant m_n, \quad (3.1.3)$$

defined, for any $x \in \mathcal{O}(n)$ and $y_1, \dots, y_n \in \mathcal{O}$, by

$$x \circ [y_1, \dots, y_n] := (\dots((x \circ_n y_n) \circ_{n-1} y_{n-1}) \dots) \circ_1 y_1. \quad (3.1.4)$$

When $\mathcal{O}$ is combinatorial as a graded set, $\mathcal{O}$ is *combinatorial*. In this case, the *Hilbert series* $\mathcal{H}_{\mathcal{O}}(t)$ of $\mathcal{O}$ is the generating series $\mathcal{G}_{\mathcal{O}}(t)$. If $\mathcal{O}_1$ and $\mathcal{O}_2$ are two operads, a map $\phi : \mathcal{O}_1 \to \mathcal{O}_2$ is an *operad morphism* if it respects arities, sends the unit of $\mathcal{O}_1$ to the unit of $\mathcal{O}_2$, and commutes with partial composition maps. We say that $\mathcal{O}_2$ is a *suboperad* of $\mathcal{O}_1$ if $\mathcal{O}_2$ is a graded subset of $\mathcal{O}_1$, $\mathcal{O}_1$ and $\mathcal{O}_2$ have the same unit, and the partial compositions of $\mathcal{O}_2$ are the ones of $\mathcal{O}_1$ restricted on $\mathcal{O}_2$. For any subset $\mathfrak{G}$ of $\mathcal{O}$, the *operad generated* by $\mathfrak{G}$ is the smallest suboperad $\mathcal{O}^{\mathfrak{G}}$ of $\mathcal{O}$ containing $\mathfrak{G}$. When $\mathcal{O}^{\mathfrak{G}} = \mathcal{O}$ and $\mathfrak{G}$ is minimal with respect to the inclusion among the subsets of $\mathfrak{G}$ satisfying this property, $\mathfrak{G}$ is a *minimal generating set* of $\mathcal{O}$ and its elements are *generators* of $\mathcal{O}$. An *operad congruence* of $\mathcal{O}$ is an equivalence relation $\equiv$ respecting the arities and such that, for any $x, y, x', y' \in \mathcal{O}$, $x \equiv x'$ and $y \equiv y'$ implies $x \circ_i y \equiv x' \circ_i y'$ for any $i \in [|x|]$. The $\equiv$-equivalence class of any $x \in \mathcal{O}$ is denoted by $[x]_{\equiv}$. Given an operad congruence $\equiv$, the *quotient operad* $\mathcal{O}/_{\equiv}$ is the operad on the set of all $\equiv$-equivalence classes and defined in the usual way.

3.2. **Presentations, rewrite relations, and bases.** We recall the notion of presentation by generators and relations of an operad. By using rewrite systems on syntax trees, this leads to the notion of bases of an operad. This notion is crucial to see the elements an operad satisfying some conditions as syntax trees factor-avoiding some patterns.

3.2.1. *Free operads and presentations.* For any graded set $\mathfrak{G}$, the *free operad* on $\mathfrak{G}$ is the operad $\mathbf{FO}(\mathfrak{G})$ wherein for any $n \geqslant 1$, $\mathbf{FO}(\mathfrak{G})(n)$ is the set $\mathbf{S}(\mathfrak{G})(n)$ of all $\mathfrak{G}$-trees of arity n. The partial compositions $\circ_i$ of $\mathbf{FO}(\mathfrak{G})$ are the partial compositions of $\mathfrak{G}$-trees (see Section 1.1.3). A *presentation* of an operad $\mathcal{O}$ is a pair $(\mathfrak{G}, \equiv)$ such that $\mathfrak{G}$ is a graded set, $\equiv$ is an operad congruence of $\mathbf{FO}(\mathfrak{G})$, and $\mathcal{O}$ is isomorphic to $\mathbf{FO}(\mathfrak{G})/_{\equiv}$. Let us also define the *evaluation map* $\mathrm{ev} : \mathbf{FO}(\mathfrak{G}) \to \mathcal{O}$ as the unique surjective operad morphism satisfying, for any $\mathsf{a} \in \mathfrak{G}$, $\mathrm{ev}(\mathfrak{c}(\mathsf{a})) = \mathsf{a}$. We call *treelike expression* on $\mathfrak{G}$ of an element x of $\mathcal{O}$ any $\mathfrak{G}$-tree of the fiber $\mathrm{ev}^{-1}(x)$.

3.2.2. *Rewrite rules on trees and pattern avoidance.* We explain here and in the next section a useful link for our purposes between presentations of operads and pattern avoidance in syntax trees. This link passes by rewrite rules on syntax trees. Notations and notions about general rewrite rules used here can be found in [BN98].

A *rewrite rule* on $\mathfrak{G}$-trees is an ordered pair $(\mathfrak{s}, \mathfrak{s}')$ of $\mathfrak{G}$-trees such that $|\mathfrak{s}| = |\mathfrak{s}'|$. A set of rewrite rules defines a binary relation $\to$ on $\mathbf{FO}(\mathfrak{G})$ for which we denote by $\mathfrak{s} \to \mathfrak{s}'$ the fact that $(\mathfrak{s}, \mathfrak{s}') \in \to$. For any set $\to$ of rewrite rules, we denote by $\Rightarrow$ the *rewrite relation induced* by $\to$ as the binary relation satisfying

$$\mathfrak{r} \circ_i \left(\mathfrak{s} \circ \left[\mathfrak{r}_1, \dots, \mathfrak{r}_{|\mathfrak{s}|}\right]\right) \Rightarrow \mathfrak{r} \circ_i \left(\mathfrak{s}' \circ \left[\mathfrak{r}_1, \dots, \mathfrak{r}_{|\mathfrak{s}|}\right]\right), \tag{3.2.1}$$

if $\mathfrak{s} \to \mathfrak{s}'$ where and $\mathfrak{r}$, $\mathfrak{r}_1$, ..., $\mathfrak{r}_{|\mathfrak{s}|}$ are $\mathfrak{G}$-trees, and $i \in [|\mathfrak{r}|]$. In other words, one has $\mathfrak{t} \Rightarrow \mathfrak{t}'$ if it is possible to obtain $\mathfrak{t}'$ from $\mathfrak{t}$ by replacing a factor $\mathfrak{s}$ of $\mathfrak{t}$ by $\mathfrak{s}'$ whenever $\mathfrak{s} \to \mathfrak{s}'$. Let also $\overset{*}{\Rightarrow}$ be the reflexive and transitive closure of $\Rightarrow$. If $\mathfrak{t}$ and $\mathfrak{t}'$ are two $\mathfrak{G}$-trees such that $\mathfrak{t} \overset{*}{\Rightarrow} \mathfrak{t}'$, then $\mathfrak{t}$ is *rewritable* into $\mathfrak{t}'$. If $\mathfrak{t}$ is a $\mathfrak{G}$-tree such that, for any $\mathfrak{G}$-tree $\mathfrak{t}'$, $\mathfrak{t} \overset{*}{\Rightarrow} \mathfrak{t}'$ implies $\mathfrak{t} = \mathfrak{t}'$, then $\mathfrak{t}'$ is a *normal form* for $\Rightarrow$. The set of all normal forms for $\Rightarrow$ is denoted by $\mathcal{N}_{\Rightarrow}$. Besides, if there is not infinite chain $\mathfrak{t}_0 \Rightarrow \mathfrak{t}_1 \Rightarrow \mathfrak{t}_2 \Rightarrow \cdots$, then $\Rightarrow$ is *terminating*. Finally, if for all $\mathfrak{G}$-trees $\mathfrak{t}$, $\mathfrak{s}_1$, and $\mathfrak{s}_2$ such that $\mathfrak{t} \overset{*}{\Rightarrow} \mathfrak{s}_1$ and $\mathfrak{t} \overset{*}{\Rightarrow} \mathfrak{s}_2$, there exists a $\mathfrak{G}$-tree $\mathfrak{t}'$ such that $\mathfrak{s}_1 \overset{*}{\Rightarrow} \mathfrak{t}'$ and $\mathfrak{s}_2 \overset{*}{\Rightarrow} \mathfrak{t}'$, then $\Rightarrow$ is *confluent*.

Let us denote by $\mathcal{P}_{\to}$ the set of the $\mathfrak{G}$-trees appearing as left members of $\to$.

Lemma 3.2.1. *If $\to$ is a set of rewrite rules on $\mathfrak{G}$-trees, then $\mathcal{N}_{\Rightarrow}$ is the set of all the $\mathfrak{G}$-trees factor-avoiding $\mathcal{P}_{\to}$.*

Proof. Assume first that $\mathfrak{t}$ is a $\mathfrak{G}$-tree factor-avoiding $\mathcal{P}_{\to}$. Then, due to the definition (3.2.1) of $\Rightarrow$, $\mathfrak{t}$ is not rewritable by $\Rightarrow$. Hence, $\mathfrak{t}$ is a normal form for $\Rightarrow$. Conversely, assume that $\mathfrak{t} \in \mathcal{N}_{\Rightarrow}$. In this case, by definition of a normal form, $\mathfrak{t}$ is not rewritable by $\Rightarrow$, so that $\mathfrak{t}$ does not admit any occurrence of a tree appearing as a left member of $\to$. □

3.2.3. *Orientations and bases.* Let $\mathcal{O}$ be an operad admitting a presentation $(\mathfrak{G}, \equiv)$. A set $\to$ of rewrite rules is an *orientation* of $\equiv$ if the reflexive, symmetric, and transitive closure of $\Rightarrow$ is $\equiv$. When $\Rightarrow$ is terminating and confluent, the orientation $\to$ of $\equiv$ is *faithful*.

Lemma 3.2.2. *Let $\mathcal{O}$ be an operad admitting a presentation $(\mathfrak{G}, \equiv)$ and $\to$ be a faithful orientation of $\equiv$. For any $n \geqslant 1$, the restriction of the evaluation map* ev *on $\mathcal{N}_{\Rightarrow}(n)$ is a bijection between this last set and $\mathcal{O}(n)$.*

Proof. Let $x \in \mathcal{O}(n)$. Since $\mathfrak{G}$ is a generating set of $\mathcal{O}$, x admits a treelike expression $\mathfrak{t}$ on $\mathfrak{G}$. Since $\Rightarrow$ is terminating, there is a $\mathfrak{G}$-tree $\mathfrak{t}' \in [\mathfrak{t}]_{\equiv}$ such that $\mathfrak{t}'$ is a normal form for $\Rightarrow$. This implies $\mathrm{ev}(\mathfrak{t}') = x$ and shows that ev is surjective.

Since $\to$ is an orientation of $\equiv$, if $\mathfrak{t}$ and $\mathfrak{t}'$ are two normal forms for $\Rightarrow$ of arity n such that $\mathrm{ev}(\mathfrak{t}) = \mathrm{ev}(\mathfrak{t}')$, then $\mathfrak{t} \equiv \mathfrak{t}'$. Since $\equiv$ is the reflexive, symmetric, and transitive closure of $\Rightarrow$, and since $\Rightarrow$ is confluent, any $\equiv$-equivalence class admits at most one normal form. Hence, $\mathfrak{t} = \mathfrak{t}'$, showing that $\Rightarrow$ is injective. □

Let $\mathcal{O}$ be an operad admitting a presentation $(\mathfrak{G}, \equiv)$. When there exists a faithful orientation $\to$ of $\equiv$, the set $\mathcal{N}_{\Rightarrow}$ is the *$\to$-basis* of $\mathcal{O}$. By Lemma 3.2.2, the is a one-to-one correspondence between the graded sets $\mathcal{N}_{\Rightarrow}$ and $\mathcal{O}$. Moreover, $\mathcal{N}_{\Rightarrow}$ can be described as the set of the trees factor-avoiding certain trees, as stated by Lemma 3.2.1. These bases was called *Poincaré-Birkhoff-Witt basis* in [Hof10] and maintain strong connections with Koszulity of operads [GK94, DK10].

3.3. **Refinements of Hilbert series and enumeration.** We introduce a refined of the Hilbert series of an operad with respect to an orientation of one of its presentations. A general strategy to count combinatorial objects with respect to their sizes and some statistics relying on operads and factor-avoidance in trees is provided.

3.3.1. *Statistics.* A *statistics* on a set X is a map $\mathrm{s} : X \to \mathbb{N}$. Let $\mathcal{O}$ be an operad admitting a presentation $(\mathfrak{G}, \equiv)$ faithfully oriented by $\rightarrow$. Let us define, for any $\mathrm{a} \in \mathfrak{G}$, the statistics s_a on $\mathcal{O}$ in the following way. For any $x \in \mathcal{O}$, we set $\mathrm{s}_\mathrm{a}(x) := \deg_\mathrm{a}(\mathfrak{t})$ where $\mathfrak{t}$ is a treelike expression on $\mathfrak{G}$ of x which is also a normal form for $\Rightarrow$. By Lemma 3.2.2, this definition is consistent since $\mathfrak{t}$ is unique among the trees satisfying these properties.

3.3.2. *Refined Hilbert series.* The *$\rightarrow$-Hilbert series* of $\mathcal{O}$ is the series $H_\rightarrow$ of $\mathbb{K}\langle\langle t, q, \mathbf{Q}_\mathfrak{G}\rangle\rangle$ defined by

$$H_\rightarrow := F\left(\mathcal{P}_\rightarrow\right). \tag{3.3.1}$$

In other words, $H_\rightarrow$ is the enumerative image of the characteristic series of the $\mathfrak{G}$-trees factor-avoiding the trees appearing as left members of $\rightarrow$.

Proposition 3.3.1. *Let $\mathcal{O}$ be a combinatorial operad admitting a presentation $(\mathfrak{G}, \equiv)$ faithfully oriented by $\rightarrow$. Then, $H_\rightarrow$ is the series wherein the coefficient of $t^n q^d q_{\mathrm{a}_1}^{\alpha_1} \dots q_{\mathrm{a}_\ell}^{\alpha_\ell}$, $n \geqslant 1$, $d \geqslant 0$, $\alpha_i \geqslant 0$, $i \in [\ell]$, is the number of elements x of $\mathcal{O}$ of arity n, degree d, and such that $\mathrm{s}_{\mathrm{a}_i}(x) = \alpha_i$ for all $i \in [\ell]$.*

Proof. By Lemmas 3.2.1 and 3.2.2, $\mathbf{F}\left(\mathcal{P}_\rightarrow\right)$ is the characteristic series of the $\rightarrow$-basis $\mathcal{N}_\Rightarrow$ of $\mathcal{O}$. The statement of the proposition follows from the definitions of the statistics s_a, $\mathrm{a} \in \mathfrak{G}$, and of the enumerative images of $\mathfrak{G}$-tree series. □

When $\mathcal{O}$ is combinatorial, observe that the $\rightarrow$-Hilbert series of $\mathcal{O}$ is a refinement of the Hilbert series of $\mathcal{O}$. Indeed, by Proposition 3.3.1, the specialization $H_{\rightarrow |q:=1, q_\mathrm{a}:=1, \mathrm{a} \in \mathfrak{G}}$ is the Hilbert series $\mathcal{H}_\mathcal{O}(t)$ of $\mathcal{O}$.

3.3.3. *Operads as tools for enumeration.* The results presented in the previous sections can be applied, together with operad theory, for enumerative prospects. Indeed, if X is a combinatorial graded set for which we want to describe its generating series $\mathcal{G}_X(t)$, a strategy consists in

(1) endowing X with partial composition maps

$$\circ_i : X(n) \times X(m) \to X(n + m - 1), \qquad 1 \leqslant i \leqslant n, 1 \leqslant m \tag{3.3.2}$$

so that X admits the structure of an operad;

(2) exhibiting a presentation $(\mathfrak{G}, \equiv)$ of the operad on X just introduced;

(3) providing a faithful orientation $\rightarrow$ of $\equiv$;

(4) computing the $\rightarrow$-Hilbert series $H_\rightarrow$ of the considered operad on X.

By Proposition 3.3.1, $H_{\to}$ is a refinement of $\mathcal{G}_X(t)$ and hence, the knowledge of $H_{\to}$ leads to the knowledge of $\mathcal{G}_X(t)$. Moreover, by Lemma 3.2.1, Proposition 2.3.4 provides a way to express $H_{\to}$ by a system of equations. Besides, this strategy to enumerate X passes by the definition of the statistics s_{a}, $\mathrm{a} \in \mathfrak{G}$, on X which can be of independent interest.

4. Examples about series from operads

This last section contains examples of application of the theory of operads for enumeration. We recall here the definitions of some operads involving combinatorial graded sets and apply the results of Sections 2 and 3 to obtain expressions for their generating series taking into account of some statistics.

To not overload the notation, the results of the previous sections are used here implicitly. Moreover, we shall not explicitly prove the faithfulness of the considered orientations. This can easily be done by using general results about rewrite rules on trees, as presented for instance in [Gir18].

4.1. **On some classical operads.**

We begin by considering some well-known and classical operads involving families of trees: bicolored Schröder trees, binary trees, and based noncrossing trees.

4.1.1. *2-associative operad.*

The *2-associative operad* [LR06] is the operad $2\mathbf{As}$ having the presentation $(\mathfrak{G}_{2\mathbf{As}}, \equiv)$ where

$$\mathfrak{G}_{2\mathbf{As}} := \mathfrak{G}_{2\mathbf{As}}(2) := \{\mathrm{a}, \mathrm{b}\}, \tag{4.1.1}$$

and $\equiv$ is the finest operad congruence satisfying

$$\mathrm{a} \circ_1 \mathrm{a} \equiv \mathrm{a} \circ_2 \mathrm{a}, \tag{4.1.2a}$$

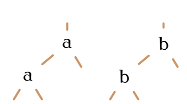
$$\mathrm{b} \circ_1 \mathrm{b} \equiv \mathrm{b} \circ_2 \mathrm{b}. \tag{4.1.2b}$$

The first dimensions of this operad are

$$1, 2, 6, 22, 90, 394, 1806, 8558 \tag{4.1.3}$$

and form Sequence **A006318** of [Slo]. This operad can be realized as an operad of bicolored Schröder trees (see for instance [Gir18]), where a *bicolored Schröder tree* is a Schröder tree such that each internal node is assigned with an element of the set $\{0, 1\}$ and all nodes that have a father labeled by 0 (resp. 1) are labeled by 1 (resp. 0). By setting that the arity of a bicolored Schröder tree is the number of its leaves, the set of all bicolored Schröder trees forms a combinatorial graded set.

The orientation $\to$ of $\equiv$ obtained by orienting (4.1.2a) and (4.1.2b) from left to right is faithful. The $\to$-Hilbert series of $2\mathbf{As}$ satisfies

$H_{\to} = F\Big($  $, \emptyset\Big)$ (4.1.4) where

$$H_{\to} = t + qq_{\mathrm{a}}F\left(\mathcal{P}_{\to}, \mathrm{a}\right) H_{\to} + qq_{\mathrm{b}}F\left(\mathcal{P}_{\to}, \mathrm{b}\right) H_{\to}, \tag{4.1.5a}$$

$$F\left(\mathcal{P}_{\to}, \mathrm{a}\right) = t + qq_{\mathrm{b}}F\left(\mathcal{P}_{\to}, \mathrm{b}\right) H_{\to}, \tag{4.1.5b}$$

$$F\left(\mathcal{P}_{\to}, \mathrm{b}\right) = t + qq_{\mathrm{a}}F\left(\mathcal{P}_{\to}, \mathrm{a}\right) H_{\to}. \tag{4.1.5c}$$

This series satisfies the algebraic equation

$$H_{\to} = \frac{t + q^2 q_{\mathrm{a}} q_{\mathrm{b}} t H_{\to}^2 + q^2 q_{\mathrm{a}} q_{\mathrm{b}} H_{\to}^3}{1 - tqq_{\mathrm{a}} - tqq_{\mathrm{b}}}, \tag{4.1.6}$$

and writes as

$$\begin{aligned} H_{\to} = t &+ (q_{\mathrm{a}} + q_{\mathrm{b}})\, qt^2 + \left(q_{\mathrm{a}}^2 + 4q_{\mathrm{a}}q_{\mathrm{b}} + q_{\mathrm{b}}^2\right) q^2t^3 + \left(q_{\mathrm{a}}^3 + 10q_{\mathrm{a}}^2q_{\mathrm{b}} + 10q_{\mathrm{a}}q_{\mathrm{b}}^2 + q_{\mathrm{b}}^3\right) q^3t^4 \\ &+ \left(q_{\mathrm{a}}^4 + 20q_{\mathrm{a}}^3q_{\mathrm{b}} + 48q_{\mathrm{a}}^2q_{\mathrm{b}}^2 + 20q_{\mathrm{a}}q_{\mathrm{b}}^3 + q_{\mathrm{b}}^4\right) q^4t^5 \\ &+ \left(q_{\mathrm{a}}^5 + 35q_{\mathrm{a}}^4q_{\mathrm{b}} + 161q_{\mathrm{a}}^3q_{\mathrm{b}}^2 + 161q_{\mathrm{a}}^2q_{\mathrm{b}}^3 + 35q_{\mathrm{a}}q_{\mathrm{b}}^4 + q_{\mathrm{b}}^5\right) q^5t^6 + \cdots . \end{aligned} \tag{4.1.7}$$

The statistics $\mathrm{s_a}$ and $\mathrm{s_b}$ are related to Triangle **A175124** of [Slo]. These statistics count the number of internal nodes labeled by 0 (or by 1) in a bicolored Schröder tree.

4.1.2. *Dipterous operad.* The *dipterous operad* [LR03] is the operad **Dipt** having the presentation $(\mathfrak{G}_{\mathbf{Dipt}}, \equiv)$ where

$$\mathfrak{G}_{\mathbf{Dipt}} := \mathfrak{G}_{\mathbf{Dipt}}(2) := \{\mathrm{a}, \mathrm{b}\}, \tag{4.1.8}$$

and $\equiv$ is the finest operad congruence satisfying

$$\mathrm{a} \circ_1 \mathrm{a} \equiv \mathrm{a} \circ_2 \mathrm{a}, \tag{4.1.9a}$$

$$\mathrm{b} \circ_1 \mathrm{b} \equiv \mathrm{b} \circ_2 \mathrm{a}. \tag{4.1.9b}$$

The dimensions of this operad are the same as the ones of 2**As** so that **Dipt** can be realized as an operad of bicolored Schröder trees.

The orientation $\to$ of $\equiv$ obtained by orienting (4.1.9a) from left to right, and (4.1.9b) from right to left is faithful. The $\to$-Hilbert series of **Dipt** satisfies

$$H_{\to} = F\left(\{\mathrm{a} \circ_1 \mathrm{a},\ \mathrm{b} \circ_2 \mathrm{a}\}, \emptyset\right) \tag{4.1.10}$$

where

$$H_{\to} = t + qq_{\mathrm{a}}F\left(\mathcal{P}_{\to}, \mathrm{a}\right) H_{\to} + qq_{\mathrm{b}}H_{\to}F\left(\mathcal{P}_{\to}, \mathrm{a}\right), \tag{4.1.11a}$$

$$F\left(\mathcal{P}_{\to}, \mathrm{a}\right) = t + qq_{\mathrm{b}}H_{\to}F\left(\mathcal{P}_{\to}, \mathrm{a}\right). \tag{4.1.11b}$$

This series satisfies the algebraic equation

$$H_{\to} = t + tqq_{\mathrm{a}}H_{\to} + qq_{\mathrm{b}}H_{\to}^2, \tag{4.1.12}$$

and writes as

$$\begin{aligned} H_{\to} = t + (q_a + q_b)\, q t^2 + \left(q_a^2 + 3q_aq_b + 2q_b^2\right) q^2 t^3 + \left(q_a^3 + 6q_a^2q_b + 10q_aq_b^2 + 5q_b^3\right) q^3t^4 \\ + \left(q_a^4 + 10q_a^3q_b + 30q_a^2q_b^2 + 35q_aq_b^3 + 14q_b^4\right) q^4t^5 \\ + \left(q_a^5 + 15q_a^4q_b + 70q_a^3q_b^2 + 140q_a^2q_b^3 + 126q_aq_b^4 + 42q_b^5\right) q^5t^6 + \cdots . \end{aligned} \tag{4.1.13}$$

The statistics s_a is related to Triangle **A060693** of [Slo], and the statistics s_b is related to Triangle **A088617** of [Slo] (one is the mirror image of the other). These statistics count the number of peaks in Schröder paths (which are some paths in one-to-one correspondence with bicolored Schröder trees).

4.1.3. *Duplicial operad.* The *duplicial operad* [Lod08] is the operad **Dup** having the presentation $(\mathfrak{G}_{\mathbf{Dup}}, \equiv)$ where

$$\mathfrak{G}_{\mathbf{Dup}} := \mathfrak{G}_{\mathbf{Dup}}(2) := \{a, b\}, \tag{4.1.14}$$

and $\equiv$ is the finest operad congruence satisfying

$$a \circ_1 a \equiv a \circ_2 a, \tag{4.1.15a}$$

$$b \circ_1 a \equiv a \circ_2 b, \tag{4.1.15b}$$

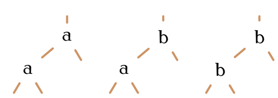
$$b \circ_1 b \equiv b \circ_2 b. \tag{4.1.15c}$$

The first dimensions of this operad are

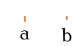
$$1, 2, 5, 14, 42, 132, 429, 1430 \tag{4.1.16}$$

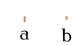
and form Sequence **A000108** of [Slo]. This operad can be realized as an operad of binary trees.

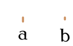
The orientation $\to$ of $\equiv$ obtained by orienting (4.1.15a), (4.1.15b), and (4.1.15c) from left to right is faithful. The $\to$-Hilbert series of **Dup** satisfies

$$H_{\to} = F\left(\right.$$  $$\left. , \emptyset \right) \tag{4.1.17}$$

where

$$H_{\to} = t + qq_aF\left(\mathscr{P}_{\to}, \right.$$ $$\left.\right) H_{\to} + qq_bF\left(\mathscr{P}_{\to}, \right.$$  $$\left.\right) H_{\to}, \tag{4.1.18a}$$

$$F\left(\mathscr{P}_{\to}, \right.$$ $$\left.\right) = t + qq_bF\left(\mathscr{P}_{\to}, \right.$$  $$\left.\right) H_{\to}, \tag{4.1.18b}$$

$$F\left(\mathscr{P}_{\to}, \right.$$  $$\left.\right) = t. \tag{4.1.18c}$$

This series satisfies the algebraic equation

$$H_{\to} = t + tqq_bH_{\to} + tqq_aH_{\to} + tq^2q_aq_bH_{\to}^2, \tag{4.1.19}$$

and writes as

$$H_{\to} = t + (q_a + q_b)\, q t^2 + \left(q_a^2 + 3q_aq_b + q_b^2\right) q^2t^3 + \left(q_a^3 + 6q_a^2q_b + 6q_aq_b^2 + q_b^3\right) q^3t^4$$

$$+\left(q_a^4 + 10q_a^3q_b + 20q_a^2q_b^2 + 10q_aq_b^3 + q_b^4\right) q^4t^5$$
$$+\left(q_a^5 + 15q_a^4q_b + 50q_a^3q_b^2 + 50q_a^2q_b^3 + 15q_aq_b^4 + q_b^5\right) q^5t^6 + \cdots . \quad (4.1.20)$$

The statistics s_a and s_b are related to Triangle **A001263** of [Slo] known as triangle of Narayana numbers [Nar55]. These statistics count the number of edges oriented to the right connecting two internal nodes in a binary tree (which are in one-to-one correspondence with the elements of **Dup**).

4.1.4. *Based noncrossing trees.* The *based noncrossing trees operad* [Cha07] (a study of algebras over this operad was provided in [Ler11]) is the operad **NCT** having the presentation $(\mathfrak{G}_{\mathbf{NCT}}, \equiv)$ where

$$\mathfrak{G}_{\mathbf{NCT}} := \mathfrak{G}_{\mathbf{NCT}}(2) := \{a, b\}, \quad (4.1.21)$$

and $\equiv$ is the finest operad congruence satisfying

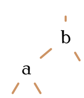
$$b \circ_1 a \equiv a \circ_2 b. \quad (4.1.22)$$

The first dimensions of this operad are

$$1, 2, 7, 30, 143, 728, 3876, 21318, 120175 \quad (4.1.23)$$

and form Sequence **A006013** of [Slo]. This operad can be realized as an operad of based noncrossing trees (see for instance [Gir18]). A *based noncrossing tree* is a configuration of chords such that the bottom side is an edge and all the edges form a tree without crossing edges. By setting that the arity of a based noncrossing tree is its number of edges, the set of all based noncrossing trees forms a combinatorial graded set.

The orientation $\rightarrow$ of $\equiv$ obtained by orienting (4.1.22) from left to right is faithful. The $\rightarrow$-Hilbert series of **NCT** satisfies

$$H_{\rightarrow} = F\left(\right.$$  $$\left. , \emptyset \right) \quad (4.1.24)$$

where

$$H_{\rightarrow} = t + qq_aH_{\rightarrow}^2 + qq_bF\left(\mathscr{P}_{\rightarrow}, \right.$$ $$\left.\right) H_{\rightarrow}, \quad (4.1.25a)$$

$$F\left(\mathscr{P}_{\rightarrow}, \right.$$ $$\left.\right) = t + qq_bF\left(\mathscr{P}_{\rightarrow}, \right.$$ $$\left.\right) H_{\rightarrow}. \quad (4.1.25b)$$

This series satisfies the algebraic equation

$$H_{\rightarrow} = t + q\left(q_a + q_b\right) H_{\rightarrow}^2 - q^2q_aq_bH_{\rightarrow}^3 = 0, \quad (4.1.26)$$

and writes as

$$H_{\rightarrow} = t + \left(q_a + q_b\right) qt^2 + \left(2q_a^2 + 3q_aq_b + 2q_b^2\right) q^2t^3 + 5\left(q_a^3 + 2q_a^2q_b + 2q_aq_b^2 + q_b^3\right) q^3t^4$$
$$+\left(14q_a^4 + 35q_a^3q_b + 45q_a^2q_b^2 + 35q_aq_b^3 + 14q_b^4\right) q^4t^5$$
$$+ 14\left(3q_a^5 + 9q_a^4q_b + 14q_a^3q_b^2 + 14q_a^2q_b^3 + 9q_aq_b^4 + 3q_b^5\right) q^5t^6 + \cdots . \quad (4.1.27)$$

The triangles related to the statistics s_a and s_b do not appear for the time being in [Slo].

4.2. On some operads from monoids. We shall consider in the following set of examples some families of combinatorial objects endowed with operad structures coming from a general construction introduced in [Gir15]. Let us recall the construction. Let $\mathcal{M}$ be a monoid, that is a set endowed with an associative product $\star$ admitting a unit $\mathbb{1}_{\mathcal{M}}$. We denote by $\mathbf{T}\mathcal{M}$ the graded set wherein for any $n \geqslant 1$, $\mathbf{T}\mathcal{M}(n)$ is the set of all words of length n on $\mathcal{M}$, seen as an alphabet. This graded set $\mathbf{T}\mathcal{M}$ is endowed with the partial composition maps $\circ_i$ defined for any $u \in \mathbf{T}\mathcal{M}(n)$, $v \in \mathbf{T}\mathcal{M}(m)$, and $i \in [n]$, by

$$u \circ_i v := u_1 \dots u_{i-1} \ (u_i \star v_1) \dots (u_i \star v_m) \ u_{i+1} \dots u_n. \tag{4.2.1}$$

It was shown in [Gir15] that $\mathbf{T}\mathcal{M}$ is an operad admitting $\mathbb{1}_{\mathcal{M}} \in \mathbf{T}\mathcal{M}(1)$ as unit. Let $\mathbb{N}$ (resp. $\mathbb{N}_\ell$) be the additive monoid of nonnegative integers (resp. the cyclic monoid of order ℓ, $\ell \geqslant 1$). In particular, the operads $\mathbf{T}\mathbb{N}$ and $\mathbf{T}\mathbb{N}_\ell$ admit suboperads whose elements can be interpreted as combinatorial objects.

4.2.1. *m-trees.* For any integer $m \geqslant 0$, an *m-tree* is a planar rooted tree wherein all internal nodes have arity $m+1$. By setting that the arity of an m-tree is its number of internal nodes, the set of all m-trees forms a combinatorial graded set.

Let $\mathbf{FCat}^{(m)}$ be the suboperad of $\mathbf{T}\mathbb{N}$ generated by the set

$$\mathfrak{G}_{\mathbf{FCat}^{(m)}} := \{00, 01, \dots, 0m\}. \tag{4.2.2}$$

It was shown in [Gir15] that there is a one-to-one correspondence between the set $\mathbf{FCat}^{(m)}(n)$ and the set of all m-trees of arity $n \geqslant 1$. Therefore, $\mathbf{FCat}^{(m)}$ is an operad on m-trees. The dimensions of this operad are provided by the Fuss-Catalan numbers so that

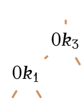
$$\#\mathbf{FCat}^{(m)}(n) = \binom{(m+1)n}{n} \frac{1}{mn+1}. \tag{4.2.3}$$

This operad admits the presentation $(\mathfrak{G}_{\mathbf{FCat}^{(m)}}, \equiv)$ where $\equiv$ is the finest operad congruence satisfying

$$\mathfrak{c}\,(0k_3) \circ_1 \mathfrak{c}\,(0k_1) \equiv \mathfrak{c}\,(0k_1) \circ_2 \mathfrak{c}\,(0k_2)\,, \qquad k_3 = k_1 + k_2. \tag{4.2.4}$$

The orientation $\to$ of $\equiv$ obtained by orienting all relations (4.2.4) from left to right is faithful. By denoting, for any $k \geqslant 0$, by $\mathbb{Q}_k$ the set $\{\mathfrak{c}\,(00)\,, \mathfrak{c}\,(01)\,, \dots, \mathfrak{c}\,(0k)\}$, the $\to$-Hilbert series of $\mathbf{FCat}^{(k)}$ satisfies

$$H_{\to} = F\left(\quad : k_1 \leqslant k_3, \emptyset \right) \tag{4.2.5}$$



where

$$H_{\to} = t + q \sum_{0 \leqslant k \leqslant m} q_{0k}\, F\left(\mathcal{P}_{\to}, \mathbb{Q}_k\right) H_{\to}, \tag{4.2.6a}$$

$$F\left(\mathcal{P}_{\to}, \mathbb{Q}_k\right) = t + q \sum_{k+1 \leqslant \ell \leqslant m} q_{0\ell}\, F\left(\mathcal{P}_{\to}, \mathbb{Q}_\ell\right) H_{\to}, \qquad 0 \leqslant k \leqslant m. \tag{4.2.6b}$$

By a straightforward computation, we obtain

$$H_{\to} = t \prod_{0 \leqslant k \leqslant m} \left(q q_{0k}\, H_{\to} + 1\right). \tag{4.2.7}$$

Let us now focus on the case $m = 1$, for which $\mathbf{FCat}^{(1)}$ is an operad on binary trees. First, as a particular case of (4.2.7), the $\rightarrow$-Hilbert series of $\mathbf{FCat}^{(1)}$ expresses as

$$\mathcal{H}_{\rightarrow} = t\left(q q_{00}\, \mathcal{H}_{\rightarrow} + 1\right)\left(q q_{01}\, \mathcal{H}_{\rightarrow} + 1\right). \tag{4.2.8}$$

This is the series (4.1.19) obtained from the operad **Dup**. Moreover, as a particular case of (4.2.4), the operad $\mathbf{FCat}^{(1)}$ admits the presentation $(\mathfrak{G}_{\mathbf{FCat}^{(1)}}, \equiv)$ where $\equiv$ is the finest operad congruence satisfying

$$\mathfrak{c}(00) \circ_1 \mathfrak{c}(00) \equiv \mathfrak{c}(00) \circ_2 \mathfrak{c}(00), \tag{4.2.9a}$$

$$\mathfrak{c}(01) \circ_1 \mathfrak{c}(00) \equiv \mathfrak{c}(00) \circ_2 \mathfrak{c}(01), \tag{4.2.9b}$$

$$\mathfrak{c}(01) \circ_1 \mathfrak{c}(01) \equiv \mathfrak{c}(01) \circ_2 \mathfrak{c}(00). \tag{4.2.9c}$$

The orientation $\rightarrow$ of $\equiv$ obtained by orienting (4.2.9a) and (4.2.9b) from left to right, and (4.2.9c) from right to left is faithful. The $\rightarrow$-Hilbert series of $\mathbf{FCat}^{(1)}$ satisfies

$$\mathcal{H}_{\rightarrow} = \mathcal{F}\left(\mathfrak{c}(00) \circ_1 \mathfrak{c}(00),\ \mathfrak{c}(01) \circ_1 \mathfrak{c}(00),\ \mathfrak{c}(01) \circ_2 \mathfrak{c}(00), \emptyset\right) \tag{4.2.10}$$

where

$$\mathcal{H}_{\rightarrow} = t + q q_{00} \mathcal{F}\left(\mathcal{P}_{\rightarrow}, \mathfrak{c}(00)\right) \mathcal{H}_{\rightarrow} + q q_{01} \mathcal{F}\left(\mathcal{P}_{\rightarrow}, \mathfrak{c}(00)\right)^2, \tag{4.2.11a}$$

$$\mathcal{F}\left(\mathcal{P}_{\rightarrow}, \mathfrak{c}(00)\right) = t + q q_{01} \mathcal{F}\left(\mathcal{P}_{\rightarrow}, \mathfrak{c}(00)\right)^2. \tag{4.2.11b}$$

This series satisfies also

$$\mathcal{H}_{\rightarrow} = \frac{1 - \sqrt{1 - 4tqq_{01}}}{q\left(q_{00}\sqrt{1 - 4tqq_{01}} - q_{00} + 2q_{01}\right)} \tag{4.2.12}$$

and writes as

$$\begin{aligned}\mathcal{H}_{\rightarrow} = t &+ (q_{00} + q_{01})\, q t^2 + \left(q_{00}^2 + 2q_{00}q_{01} + 2q_{01}^2\right) q^2 t^3 + \left(q_{00}^3 + 3q_{00}^2 q_{01} + 5q_{00}q_{01}^2 + 5q_{01}^3\right) q^3 t^4 \\ &+ \left(q_{00}^4 + 4q_{00}^3 q_{01} + 9q_{00}^2 q_{01}^2 + 14q_{00}q_{01}^3 + 14q_{01}^4\right) q^4 t^5 \\ &+ \left(q_{00}^5 + 5q_{00}^4 q_{01} + 14q_{00}^3 q_{01}^2 + 28q_{00}^2 q_{01}^3 + 42q_{00}q_{01}^4 + 42q_{01}^5\right) q^5 t^6 + \cdots,\end{aligned} \tag{4.2.13}$$

The statistics s_{00} and s_{01} are related to Triangles **A033184** and **A009766** of [Slo], known as (mirror image of) Catalan triangle. These statistics count the jump-length in a binary tree (see for instance [Kra04]).

4.2.2. *Schröder trees.* A *Schröder tree* is a planar rooted tree wherein all internal nodes have arity 2 or more. By setting that the arity of a Schröder tree is its number of leaves minus 1, the set of all Schröder trees forms a combinatorial graded set.

Let **Schr** be the suboperad of $\mathbf{T}\mathbb{N}$ generated by the set

$$\mathfrak{G}_{\mathbf{Schr}} := \{00, 01, 10\}. \tag{4.2.14}$$

It was shown in [Gir15] that there is a one-to-one correspondence between the set $\mathbf{Schr}(n)$ and the set of all Schröder trees of arity $n \geqslant 1$. Therefore, **Schr** is an operad on Schröder trees. The first dimensions of this operad are

$$1, 3, 11, 45, 197, 903, 4279, 20793 \tag{4.2.15}$$

and form Sequence A001003 of [Slo]. This operad admits the presentation $(\mathfrak{G}_{\mathbf{Schr}}, \equiv)$ where $\equiv$ is the finest operad congruence satisfying

$$\mathfrak{c}\,(00) \circ_1 \mathfrak{c}\,(00) \equiv \mathfrak{c}\,(00) \circ_2 \mathfrak{c}\,(00)\,, \tag{4.2.16a}$$

$$\mathfrak{c}\,(01) \circ_1 \mathfrak{c}\,(10) \equiv \mathfrak{c}\,(10) \circ_2 \mathfrak{c}\,(01)\,, \tag{4.2.16b}$$

$$\mathfrak{c}\,(00) \circ_1 \mathfrak{c}\,(01) \equiv \mathfrak{c}\,(00) \circ_2 \mathfrak{c}\,(10)\,, \tag{4.2.16c}$$

$$\mathfrak{c}\,(01) \circ_1 \mathfrak{c}\,(00) \equiv \mathfrak{c}\,(00) \circ_2 \mathfrak{c}\,(01)\,, \tag{4.2.16d}$$

$$\mathfrak{c}\,(00) \circ_1 \mathfrak{c}\,(10) \equiv \mathfrak{c}\,(10) \circ_2 \mathfrak{c}\,(00)\,, \tag{4.2.16e}$$

$$\mathfrak{c}\,(01) \circ_1 \mathfrak{c}\,(01) \equiv \mathfrak{c}\,(01) \circ_2 \mathfrak{c}\,(00)\,, \tag{4.2.16f}$$

$$\mathfrak{c}\,(10) \circ_1 \mathfrak{c}\,(00) \equiv \mathfrak{c}\,(10) \circ_2 \mathfrak{c}\,(10)\,. \tag{4.2.16g}$$

The orientation $\to$ of $\equiv$ obtained by orienting (4.2.16a), (4.2.16b), (4.2.16c), (4.2.16d), (4.2.16e), and (4.2.16f) from left to right, and (4.2.16g) from right to left is faithful. The $\to$-Hilbert series of **Schr** satisfies

$$H_{\to} = F\left(\begin{smallmatrix} & 00 \\ 00 & \end{smallmatrix}\; \begin{smallmatrix} & 01 \\ 10 & \end{smallmatrix}\; \begin{smallmatrix} & 00 \\ 01 & \end{smallmatrix}\; \begin{smallmatrix} & 01 \\ 00 & \end{smallmatrix}\; \begin{smallmatrix} & 00 \\ 10 & \end{smallmatrix}\; \begin{smallmatrix} & 01 \\ 01 & \end{smallmatrix}\; \begin{smallmatrix} 10 & \\ & 10 \end{smallmatrix}\;, \emptyset \right) \tag{4.2.17}$$

where

$$H_{\to} = t + q q_{00} F\left(\mathscr{P}_{\to},\; 00 \;\; 01 \;\; 10\right) H_{\to} + q q_{01} F\left(\mathscr{P}_{\to},\; 00 \;\; 01 \;\; 10\right) H_{\to} + q q_{10} H_{\to} F\left(\mathscr{P}_{\to},\; 10\right), \tag{4.2.18a}$$

$$F\left(\mathscr{P}_{\to},\; 00 \;\; 01 \;\; 10\right) = t, \tag{4.2.18b}$$

$$F\left(\mathscr{P}_{\to},\; 10\right) = t + q q_{00} F\left(\mathscr{P}_{\to},\; 00 \;\; 01 \;\; 10\right) H_{\to} + q q_{01} F\left(\mathscr{P}_{\to},\; 00 \;\; 01 \;\; 10\right) H_{\to}. \tag{4.2.18c}$$

This series satisfies the algebraic equation

$$t + \left(tq\left(q_{00} + q_{01} + q_{10}\right) - 1\right) H_{\to} + \left(tq^2\left(q_{00}q_{10} + q_{01}q_{10}\right)\right) H_{\to}^2 = 0 \tag{4.2.19}$$

and writes as

$$H_{\to} = t + \left(q_{00} + q_{01} + q_{10}\right) q t^2 + \left(q_{00}^2 + 2q_{00}q_{01} + 3q_{00}q_{10} + q_{01}^2 + 3q_{01}q_{10} + q_{10}^2\right) q^2 t^3$$
$$+ \left(q_{00}^3 + 3q_{00}^2 q_{01} + 6q_{00}^2 q_{10} + 3q_{00}q_{01}^2 + 12q_{00}q_{01}q_{10} + 6q_{00}q_{10}^2\right.$$

$$+q_{01}^3 + 6q_{01}^2 q_{10} + 6q_{01} q_{10}^2 + q_{10}^3\big) q^3 t^4 + \cdots . \quad (4.2.20)$$

The statistics s_{00} and s_{10} are related to Triangle **A126216** of [Slo], and the statistics s_{01} is related to Triangle **A114656** of [Slo].

4.2.3. *Motzkin paths.* A *Motzkin path* is a path in $\mathbb{N}^2$ connecting the points $(0, 0)$ and $(n - 1, 0)$ by steps in the set $\{(1, -1), (1, 0), (1, 1)\}$. By setting that the arity of a Motzkin path is n, the set of all Motzkin paths forms a combinatorial graded set.

Let **Motz** be the suboperad of $\mathbf{T}\mathbb{N}$ generated by the set

$$\mathfrak{G}_{\mathbf{Motz}} := \{00, 010\} . \quad (4.2.21)$$

It was shown in [Gir15] that there is a one-to-one correspondence between the set $\mathbf{Motz}(n)$ and the set of all Motzkin paths of arity $n \geqslant 1$. Therefore, **Motz** is an operad on Motzkin paths. The first dimensions of this operad are

$$1, 1, 2, 4, 9, 21, 51, 127 \quad (4.2.22)$$

and form Sequence **A001006** of [Slo]. This operad admits the presentation $(\mathfrak{G}_{\mathbf{Motz}}, \equiv)$ where $\equiv$ is the finest operad congruence satisfying

$$\mathfrak{c}(00) \circ_1 \mathfrak{c}(00) \equiv \mathfrak{c}(00) \circ_2 \mathfrak{c}(00) , \quad (4.2.23a)$$

$$\mathfrak{c}(010) \circ_1 \mathfrak{c}(00) \equiv \mathfrak{c}(00) \circ_2 \mathfrak{c}(010) , \quad (4.2.23b)$$

$$\mathfrak{c}(00) \circ_1 \mathfrak{c}(010) \equiv \mathfrak{c}(010) \circ_3 \mathfrak{c}(00) , \quad (4.2.23c)$$

$$\mathfrak{c}(010) \circ_1 \mathfrak{c}(010) \equiv \mathfrak{c}(010) \circ_3 \mathfrak{c}(010) . \quad (4.2.23d)$$

The orientation $\to$ of $\equiv$ obtained by orienting (4.2.23a), (4.2.23b), (4.2.23c), and (4.2.23d) from left to right is faithful. The $\to$-Hilbert series of **Motz** satisfies

$$H_{\to} = F\left(\mathfrak{c}(00) \circ_1 \mathfrak{c}(00),\ \mathfrak{c}(010) \circ_1 \mathfrak{c}(00),\ \mathfrak{c}(00) \circ_1 \mathfrak{c}(010),\ \mathfrak{c}(010) \circ_1 \mathfrak{c}(010), \emptyset\right) \quad (4.2.24)$$

where

$$H_{\to} = t + qq_{00} F\left(\mathcal{P}_{\to}, \mathfrak{c}(00)\ \mathfrak{c}(010)\right) H_{\to} + qq_{010} F\left(\mathcal{P}_{\to}, \mathfrak{c}(00)\ \mathfrak{c}(010)\right) H_{\to}^2, \quad (4.2.25a)$$

$$F\left(\mathcal{P}_{\to}, \mathfrak{c}(00)\ \mathfrak{c}(010)\right) = t. \quad (4.2.25b)$$

This series satisfies the algebraic equation

$$H_{\to} = t + tqq_{00} H_{\to} + tqq_{010} H_{\to}^2 \quad (4.2.26)$$

and writes as

$$H_{\to} = t + qq_{00}t^2 + \left(q^2 q_{00}^2 + qq_{010}\right) t^3 + \left(q^3 q_{00}^3 + 3q^2 q_{00} q_{010}\right) t^4 + \left(q^4 q_{00}^4 + 6q^3 q_{00}^2 q_{010} + 2q^2 q_{010}^2\right) t^5 + \left(q^5 q_{00}^5 + 10q^4 q_{00}^3 q_{010} + 10q^3 q_{00} q_{010}^2\right) t^6 + \left(q^6 q_{00}^6 + 15q^5 q_{00}^4 q_{010} + 30q^4 q_{00}^2 q_{010}^2 + 5q^3 q_{010}^3\right) t^7 + \cdots . \quad (4.2.27)$$

The statistics s_{00} and s_{010} are related to Triangle **A055151** of [Slo]. These statistics count the number of steps $(1,1)$ in a Motzkin path.

4.2.4. *Directed animals.* A *directed animal* is a finite subset A of $\mathbb{N}^2$ containing $(0,0)$ and if $(x,y) \in A \setminus \{(0,0)\}$, then $(x-1,y) \in A$ or $(x,y-1) \in A$. By setting that the arity of a directed animal is its cardinality, the set of all directed animals forms a combinatorial graded set.

Let $\mathbf{DA}$ be the suboperad of $\mathbf{T}\mathbb{N}_3$ generated by the set

$$\mathfrak{G}_{\mathbf{DA}} := \{00, 01\}. \tag{4.2.28}$$

It was shown in [Gir15] that there is a one-to-one correspondence between the set $\mathbf{DA}(n)$ and the set of all directed animals of arity $n \geqslant 1$. Therefore, $\mathbf{DA}$ is an operad on directed animals. The first dimensions of this operad are

$$1, 2, 5, 13, 35, 96, 267, 750 \tag{4.2.29}$$

and form Sequence **A005773** of [Slo]. This operad admit the presentation $(\mathfrak{G}_{\mathbf{DA}}, \equiv)$ where $\equiv$ is the finest operad congruence satisfying

$$\mathfrak{c}(00) \circ_1 \mathfrak{c}(00) \equiv \mathfrak{c}(00) \circ_2 \mathfrak{c}(00), \tag{4.2.30a}$$

$$\mathfrak{c}(01) \circ_1 \mathfrak{c}(00) \equiv \mathfrak{c}(00) \circ_2 \mathfrak{c}(01), \tag{4.2.30b}$$

$$\mathfrak{c}(01) \circ_1 \mathfrak{c}(01) \equiv \mathfrak{c}(01) \circ_2 \mathfrak{c}(00), \tag{4.2.30c}$$

$$(\mathfrak{c}(00) \circ_1 \mathfrak{c}(01)) \circ_2 \mathfrak{c}(01) \equiv (\mathfrak{c}(01) \circ_2 \mathfrak{c}(01)) \circ_3 \mathfrak{c}(01). \tag{4.2.30d}$$

The orientation $\to$ of $\equiv$ obtained by orienting (4.2.30a), (4.2.30b) from left to right, and (4.2.30c) and (4.2.30d) from right to left, is faithful. The $\to$-Hilbert series of $\mathbf{DA}$ satisfies

$$\mathcal{H}_{\to} = F\left(\mathfrak{c}(00)\circ_1\mathfrak{c}(00),\ \mathfrak{c}(01)\circ_1\mathfrak{c}(00),\ \mathfrak{c}(01)\circ_2\mathfrak{c}(00),\ (\mathfrak{c}(01)\circ_2\mathfrak{c}(01))\circ_3\mathfrak{c}(01), \emptyset\right) \tag{4.2.31}$$

where

$$\mathcal{H}_{\to} = t + qq_{00} F\left(\mathcal{P}_{\to}, \mathfrak{c}(00)\right) \mathcal{H}_{\to} + qq_{01} F\left(\mathcal{P}_{\to}, \mathfrak{c}(00)\right) F\left(\mathcal{P}_{\to}, \mathfrak{c}(00), \mathfrak{c}(01)\circ_2\mathfrak{c}(01)\right), \tag{4.2.32a}$$

$$F\left(\mathcal{P}_{\to}, \mathfrak{c}(00)\right) = t + qq_{01} F\left(\mathcal{P}_{\to}, \mathfrak{c}(00)\right) F\left(\mathcal{P}_{\to}, \mathfrak{c}(00), \mathfrak{c}(01)\circ_2\mathfrak{c}(01)\right), \tag{4.2.32b}$$

$$F\left(\mathcal{P}_{\to}, \mathfrak{c}(00), \mathfrak{c}(01)\circ_2\mathfrak{c}(01)\right) = t + qq_{01} F\left(\mathcal{P}_{\to}, \mathfrak{c}(00)\right) F\left(\mathcal{P}_{\to}, \mathfrak{c}(00), \mathfrak{c}(01), \mathfrak{c}(01)\circ_2\mathfrak{c}(01)\right), \tag{4.2.32c}$$

$$F\left(\mathcal{P}_{\to}, \mathfrak{c}(00), \mathfrak{c}(01), \mathfrak{c}(01)\circ_2\mathfrak{c}(01)\right) = t. \tag{4.2.32d}$$

This series satisfies

$$H_{\to} = \frac{1 - \sqrt{1 - 2tqq_{01} - 3t^2q^2q_{01}^2} - tq(2q_{00} + q_{01})}{2tq^2\left(q_{00}^2 + q_{00}q_{01} + q_{01}^2\right) - 2qq_{00}} \tag{4.2.33}$$

and writes as

$$\begin{aligned} H_{\to} = t &+ (q_{00} + q_{01})\, q t^2 + \left(q_{00}^2 + 2q_{00}q_{01} + 2q_{01}^2\right) q^2 t^3 + \left(q_{00}^3 + 3q_{00}^2 q_{01} + 5q_{00}q_{01}^2 + 4q_{01}^3\right) q^3 t^4 \\ &+ \left(q_{00}^4 + 4q_{00}^3 q_{01} + 9q_{00}^2 q_{01}^2 + 12 q_{00} q_{01}^3 + 9q_{01}^4\right) q^4 t^5 \\ &+ \left(q_{00}^5 + 5q_{00}^4 q_{01} + 14 q_{00}^3 q_{01}^2 + 25 q_{00}^2 q_{01}^3 + 30 q_{00} q_{01}^4 + 21 q_{01}^5\right) q^5 t^6 + \cdots , \end{aligned} \tag{4.2.34}$$

The statistics s_{00} is related to Triangle **A064189** of [Slo], and the statistics s_{01} is related to Triangle **A026300** of [Slo] (one is the mirror image of the other).

References

[BN98] F. Baader and T. Nipkow. *Term rewriting and all that*. Cambridge University Press, Cambridge, 1998. 19

[Cha02] F. Chapoton. Rooted trees and an exponential-like series. `arXiv:math/0209104`, 2002. 2

[Cha07] F. Chapoton. The anticyclic operad of moulds. *Int. Math. Res. Notices*, 20:Art. ID rnm078, 36, 2007. 25

[Cha08] F. Chapoton. Operads and algebraic combinatorics of trees. *Sém. Lothar. Combin.*, 58, 2008. 2

[DK10] V. Dotsenko and A. Khoroshkin. Gröbner bases for operads. *Duke Math. J.*, 153(2):363–396, 2010. 2, 20

[Gir15] S. Giraudo. Combinatorial operads from monoids. *J. Algebr. Comb.*, 41(2):493–538, 2015. 26, 28, 29, 30

[Gir18] S. Giraudo. *Nonsymmetric Operads in Combinatorics*. Springer Nature Switzerland AG, 2018. 2, 3, 18, 22, 25

[Gir19] S. Giraudo. Colored operads, series on colored operads, and combinatorial generating systems. *Discrete Math.*, 2019. To appear. 2

[GJ79] I. P. Goulden and D. M. Jackson. An inversion theorem for cluster decompositions of sequences with distinguished subsequences. *J. London Math. Soc.*, 20(3):567–576, 1979. 14

[GK94] V. Ginzburg and M. M. Kapranov. Koszul duality for operads. *Duke Math. J.*, 76(1):203–272, 1994. 2, 20

[Hof10] E. Hoffbeck. A Poincaré-Birkhoff-Witt criterion for Koszul operads. *Manuscripta Math.*, 131(1-2):87–110, 2010. 2, 20

[Knu97] D. Knuth. *The Art of Computer Programming, volume 1: Fundamental Algorithms*. Addison Wesley Longman, 3rd edition, 1997. 3

[KP15] A. Khoroshkin and D. Piontkovski. On generating series of finitely presented operads. *J. Algebra*, 426:377–429, 2015. 2, 17

[Kra04] W. Krandick. Trees and jumps and real roots. *J. Comput. Appl. Math.*, 162(1):51–55, 2004. 27

[Ler11] P. Leroux. L-algebras, triplicial-algebras, within an equivalence of categories motivated by graphs. *Comm. Algebra*, 39(8):2661–2689, 2011. 25

[Lod05] J.-L. Loday. Inversion of integral series enumerating planar trees. *Sém. Lothar. Combin.*, 53, 2005. Art. B53d. 2, 14

[Lod08] J.-L. Loday. Generalized bialgebras and triples of operads. *Astérisque*, 320:x+116, 2008. 24

[LR03] J.-L. Loday and M. Ronco. Algèbres de Hopf colibres. *C. R. Math.*, 337(3):153–158, 2003. 23

[LR06] J.-L. Loday and M. Ronco. On the structure of cofree Hopf algebras. *J. Reine Angew. Math.*, 592:123–155, 2006. 22

[LV12] J.-L. Loday and B. Vallette. *Algebraic Operads*, volume 346 of *Grundlehren der mathematischen Wissenschaften*. Springer, Heidelberg, 2012. Pages xxiv+634. 2

[Mé15] M. A. Méndez. *Set operads in combinatorics and computer science*. SpringerBriefs in Mathematics. Springer, Cham, 2015. 2

[Nar55] T.V. Narayana. Sur les treillis formés par les partitions d'un entier et leurs applications à la théorie des probabilités. *C. R. Acad. Sci. Paris*, 240:1188–1189, 1955. 25

[NZ99] J. Noonan and D. Zeilberger. The Goulden-Jackson cluster method: extensions, applications and implementations. *J. Differ. Equ. Appl.*, 5(4-5):355–377, 1999. 14

[Par93] S. F. Parker. *The combinatorics of functional composition and inversion*. PhD thesis, Brandeis University, 1993. 2, 14

[Row10] E. S. Rowland. Pattern avoidance in binary trees. *J. Comb. Theory A*, 117(6):741–758, 2010. 2

[Slo] N. J. A. Sloane. The On-Line Encyclopedia of Integer Sequences. `https://oeis.org/`. 22, 23, 24, 25, 27, 28, 29, 30, 31

Université Paris-Est, LIGM (UMR 8049), CNRS, ENPC, ESIEE Paris, UPEM, F-77454, Marne-la-Vallée, France

E-mail address: `samuele.giraudo@u-pem.fr`